

Gestandaardiseerde toetsen voor gelijke onderwijskansen

Een mixed-method onderzoek naar de
invloed van netgebonden toetsen op het
getuigschrift basisonderwijs

Stijn Kerkhofs

s0196051

Masterproef

Master in de Opleidings- en Onderwijswetenschappen

Promotor

dr. Roos Van Gasse

Medebeoordelaar

dr. Wanda Sass



Deze masterproef is een examendocument dat niet werd gecorrigeerd voor eventueel vastgestelde fouten. In publicaties mag naar dit werk worden gerefereerd, mits schriftelijke toelating van de promotor(en) vermeld op deze titelpagina.

Universiteit Antwerpen

| Faculteit Sociale Wetenschappen

Dankwoord

Met het schrijven van dit dankwoord, één van de laatste wapenfeiten van deze masterproef, komt het besef dat de vier OOW-jaren er stilaan op zitten. Het was een enorm leerrijke ervaring waarvan deze masterproef het sluitstuk is.

Ik wil mijn promotor, dr. Roos Van Gasse, graag bedanken voor haar feedback en tips tijdens het masterproeftraject. Hierdoor kon ik mijn masterproef telkens verbeteren, en is het eindproduct geworden wat het nu is.

Bij uitbreiding wil ik iedereen bedanken die heeft bijgedragen tot het vorm krijgen van de masterproef. Bedankt aan iedereen die de survey heeft ingevuld, een interview heeft gegeven, respondenten heeft geronseld of stukken heeft nagelezen.

Ten slotte wil ik graag mijn vrienden en familie bedanken voor hun steun de voorbije vier jaar, en voor het begrip voor mijn beperktere sociale leven tijdens die periode. Zeker tegenover mijn vriendin wil ik mij hierbij verontschuldigen voor mijn kluzenaarsbestaan en haar bedanken voor de manier waarop ze hier mee omging.

Toelichting van aanpak en eigen inbreng

Samenvattingen

Nederlands

Dit onderzoek gaat na hoe het gebruik van data uit gestandaardiseerde toetsen kan helpen bij het bevorderen van gelijke onderwijskansen voor leerlingen met een lage sociaal-economische status (SES). Er is in de literatuur geen eenduidigheid over de mogelijke invloed van gestandaardiseerde toetsen op gelijke onderwijskansen. Het huidige onderzoek richt zich specifiek op de manier waarop leerkrachten de data uit netgebonden toetsen gebruiken bij de beslissing over het toekennen van een getuigschrift basisonderwijs, en de impact hiervan op gelijke onderwijskansen.

Ten eerste wordt nagegaan in welke mate er een verband is tussen het gebruik van netgebonden toetsen en de perceptie van leerkrachten tegenover GOK-leerlingen. Ten tweede brengen we in kaart op welke manier leerkrachten basisonderwijs gebruik maken van de resultaten van netgebonden toetsen op leerlingniveau bij de beslissing over het getuigschrift. Het onderzoek maakt gebruik van een mixed-method design waarbij kwantitatieve data uit een survey en kwalitatieve data uit interviews afzonderlijk worden verzameld, en daarna met elkaar worden gecombineerd.

Uit de resultaten blijkt dat leraren zich vooral baseren op de eigen toetsen en de netgebonden toetsen eerder aanvullend gebruiken. Er werd geen effect gevonden van de mate waarin leerkrachten de netgebonden toetsen gebruiken op hun percepties tegenover GOK-leerlingen. Naast de data uit de eigen en netgebonden toetsen, gaan leraren ook grotendeels af op hun intuïtie bij het nemen van de beslissing over het getuigschrift. Dit laat ruimte om kenmerken die leerlingen met een lagere SES benadelen zoals motivatie, werkhouding, en ouderbetrokkenheid mee te nemen in de beslissing. De conclusies van dit onderzoek kunnen bijdragen aan het verder vormgeven van gestandaardiseerde toetsen in Vlaanderen, om zo gelijke onderwijskansen te bevorderen.

English

This study examines how the use of data from standardized tests can help promote equal educational opportunity for students with low socioeconomic status (SES). There is no agreement in the literature on the potential impact of standardized tests on educational equity. The current research focuses specifically on how teachers use data from net-based tests when deciding whether to award a certificate in elementary education, and its impact on equal educational opportunity.

First, we examine the extent to which there is a relationship between the use of net-based tests and teachers' perception toward GOK students. Second, we identify the ways in which elementary school teachers use student-level results from net-based tests when deciding on the certificate. The study uses a mixed-method design in which quantitative data from a survey and qualitative data from interviews are collected separately, and then combined.

The results show that teachers rely primarily on their own tests and are more likely to use the net-based tests additionally. No effect was found of the extent to which teachers use the net-based tests on their perceptions toward GOK students. In addition to the data from the own and net-based tests, teachers also largely rely on their intuition when making decisions about the certificate. This leaves room for characteristics that disadvantage lower SES students such as motivation, work ethic, and parent involvement to be factored into the decision. The conclusions of this study can contribute to further shaping standardized tests in Flanders to promote equal educational opportunities.

Blogpost

Inhoud

1. Inleiding en probleemstelling	9
1.1. Relevantie en situering.....	9
1.2. Vorig onderzoek en kennisprobleem	10
1.3. Huidig onderzoek	10
2. Theoretisch kader	12
2.1. Beslissingsprocessen	12
2.1.1. Kader voor datagebaseerde beslissingen.....	12
2.1.2. Getuigschrift basisonderwijs	12
2.2. Percepties tegenover GOK-leerlingen	13
2.2.1. Invloed van SES op studieadvies	13
2.2.2. GOK-indicatoren.....	14
2.3. Betrouwbaarheid en validiteit van de beschikbare informatie.....	15
2.3.1. Valkuilen bij evaluaties door leerkrachten.....	15
2.3.2. Kansen bij het gebruik van gestandaardiseerde toetsen	15
2.3.3. Valkuilen bij het gebruik van gestandaardiseerde toetsen.....	16
2.3.4. Complementaire vormen van beoordelen.....	16
2.4. Gebruik van netgebonden toetsen	16
2.4.1. Situering van de netgebonden toetsen.....	17
2.4.2. Gebruik op school- en leerlingniveau.....	17
3. Methodologie	18
3.1. Onderzoeksdesign	18
3.2. Kwantitatief luik (OV1)	18
3.2.1. Respondenten	18
3.2.2. Dataverzameling.....	18
3.2.3. Analyse	20
3.3. Kwalitatief luik (OV2).....	20
3.3.1. Respondenten	20
3.3.2. Dataverzameling.....	21
3.3.3. Analyse	22
4. Onderzoeksresultaten	23
4.1. Onderzoeksvraag 1.....	23
4.1.1. Beschrijvende statistiek.....	23
4.1.2. Verbanden tussen de variabelen.....	24
4.2. Onderzoeksvraag 2.....	27
4.2.1. Achtergrondkenmerken van de leerlingen	27
4.2.2. Twijfels over het getuigschrift	27

4.2.3.	Voor- en nadelen van de verschillende informatiebronnen	29
4.2.4.	Adviezen bij de beslissing	30
5.	Discussie en conclusie	31
5.1.	Bevindingen uit het huidige onderzoek	31
5.1.1.	Rationele en intuïtieve processen	31
5.1.2.	Biases tegenover GOK-leerlingen	32
5.1.3.	Implicaties voor de praktijk	32
5.2.	Beperkingen van het huidige onderzoek	33
5.3.	Suggesties voor vervolgonderzoek	33
6.	Referentielijst	34
7.	Bijlagen	39
7.1.	Bijlage 1: Verklaring op eer	39
7.2.	Bijlage 2: Survey	40
7.3.	Bijlage 3: Interviewleidraad	43
7.4.	Bijlage 4: Voorbeeld van een transcript	45
7.5.	Bijlage 5: Codeboom	46

1. Inleiding en probleemstelling

1.1. Relevantie en situering

De PISA-resultaten van 2018 tonen aan dat er in Vlaanderen een hoge mate van onderwijsongelijkheid heerst. Sociaal-economische status (SES) is een belangrijke voorspeller van zowel cognitieve als niet-cognitieve uitkomsten (Franck & Nicaise, 2018). Dit wijst op een grote mate van segregatie in het Vlaamse onderwijssysteem. Andere vormen van veelvoorkomende segregatie in Vlaanderen zijn het watervalstelsel, het aantal zittenblijvers en het hoge aantal leerlingen in het buitengewoon onderwijs (Franck & Nicaise, 2018; Valcke & Standaert, 2018).

Een voorbeeld van deze ongelijke onderwijskansen is de oriëntering van leerlingen in het secundair onderwijs. Leerlingen met een lagere SES worden sneller georiënteerd naar niet-cognitieve studierichtingen dan leerlingen met een hogere SES, ook wanneer er gecorrigeerd wordt voor eerdere resultaten (Boone & Van Houtte, 2013; Timmermans et al., 2015; van Leest et al., 2021). Stereotypen van leerkrachten over leerlingen met een lagere SES, bijvoorbeeld dat zij minder gemotiveerd of ambitieus zouden zijn, dragen bij aan deze ongelijkheid (Dunkake & Schuchart, 2015; Geven et al., 2018; Lorenz, 2021). In het huidige onderzoek kijken we specifiek naar de invloed van SES op het al dan niet toekennen van het getuigschrift basisonderwijs.

Om deze negatieve impact van SES te keren, werd in 2002 het Decreet op de Gelijke Onderwijskansen (GOK) goedgekeurd. Dit decreet zorgde voor een recht op inschrijving en dus een verplichting voor de scholen om elke leerling die voldoet aan de voorwaarden, toe te laten. Wanneer het aantal 'GOK-leerlingen' in een school meer dan 10% is, ontvangt de school extra lestijden en werkingsmiddelen. Sinds 2012 worden de thuistaal, het opleidingsniveau van de moeder en het ontvangen van een schooltoelage of studiebeurs bekeken als indicatoren van SES (Valcke & Standaert, 2018). Het GOK-beleid in Vlaanderen heeft volgens Franck et al. (2017) de sociale ongelijkheid in het onderwijs onvoldoende verminderd. Hoewel er enige verbetering zichtbaar was in PISA-resultaten op leerlingniveau tussen 2003 en 2015, met name in wiskunde- en leesprestaties en de afname van SES-invloed en genderongelijkheid, bleef de samenhang tussen school-SES en cognitieve prestaties gelijk. Bovendien nam de segregatie toe.

Uit eerder onderzoek blijkt dat een overheid kan inzetten op meer kansengelijkheid door meer standaardisering in te bouwen in het onderwijsstelsel (Geven et al., 2018; Mijs & van de Werfhorst, 2016). Standaardisering betekent dat het schoolsysteem in hoge mate gebaseerd is op een centraal vastgelegde standaard. Deze standaardisering kan zowel op vlak van input als output. Standaardisering van input kan gaan van de algemene leerdoelen tot de letterlijke tekst van het lesmateriaal. Een voorbeeld van standaardisering van output is een centraal examen.

In Vlaanderen wordt de grondwettelijke vrijheid van onderwijs hoog in het vaandel gedragen. De scholen, en bij uitbreiding de onderwijskoepels, hebben een grote autonomie in hoe zij hun pedagogisch project willen vormgeven. Sinds 1997 is er met de eindtermen al een zekere standaardisering van input gekomen. Standaardisering van output, bijvoorbeeld door een centraal examen, werd vooralsnog afgehouden (Penninckx et al., 2017; Valcke & Standaert, 2018). De onderwijskoepels maakten van de onderwijsvrijheid gebruik door zelf gestandaardiseerde toetsen te ontwikkelen voor het eigen net. Een gevolg hiervan is dat de inhoud van de toetsen gebaseerd is op de eigen leerplannen in plaats van de eindtermen, en dat vergelijkende analyses op macroniveau niet mogelijk zijn (Penninckx et al., 2017).

1.2. Vorig onderzoek en kennisprobleem

Eerder onderzoek toonde aan dat standaardisering ervoor kan zorgen dat de invloed van het ouderlijke milieu op de schoolprestaties kleiner wordt, wat een positief effect heeft op de kansengelijkheid (Mijs & van de Werfhorst, 2016). Ook de invloed van het subjectieve oordeel van de leerkracht kan afnemen wanneer beslissingen gebaseerd zijn op het resultaat van gestandaardiseerde toetsen (Vanlommel, 2018; Wyatt-Smith & Klenowski, 2013). Op die manier zou de invloed van stereotypen over leerlingen met een lagere SES verkleind kunnen worden. Wanneer leerkrachten de oriëntering van leerlingen baseren op gestandaardiseerde toetsen in plaats van het eigen aanvoelen, heeft de SES van de leerling geen direct effect meer op de beslissing (van Leest et al., 2021).

Er zijn echter ook tegenstanders van het gebruik van gestandaardiseerde toetsen. Zij geven aan dat het gebruik ervan ongelijkheid net in de hand werkt. Gestandaardiseerde toetsen kunnen leerlingen met een andere socio-culturele achtergrond of lagere SES benadelen doordat de toetsvragen doorgaans minder zijn afgestemd op hun culturele kapitaal (Au, 2015; Rogers-Ard et al., 2013). Daarnaast zijn gestandaardiseerde toetsen vaak momentopnames die maar een beperkt beeld geven van de capaciteiten van een leerling. Om tot valide beslissingen te komen, kan het daarom zinvol zijn dat leerkrachten deze data combineren met andere bronnen zoals toetsen in de klas en contextinformatie die bepaalde resultaten kunnen verklaren (Allal, 2013; Klenowski, 2014).

Er bestaat nog veel onduidelijkheid over de invloed van gestandaardiseerde toetsen op de gelijke onderwijskansen van leerlingen. Onderzoek naar het gebruik van gestandaardiseerde toetsen in de Vlaamse context richt zich slechts in beperkte mate op het creëren van gelijke onderwijskansen maar eerder op het verhogen van de onderwijskwaliteit in het algemeen en het stimuleren van de interne kwaliteitszorg van scholen. Er is daarnaast tot op heden weinig gelijkgezindheid over de invloed van gestandaardiseerde toetsen op de gelijke onderwijskansen van leerlingen. Sommige onderzoekers benoemen de objectiviteit van de gestandaardiseerde toetsen als een voordeel. Anderen vinden deze subjectieve contextinformatie net belangrijk voor de validiteit van de beslissing. Er is nood aan meer zicht op de mate waarin gestandaardiseerde toetsen kunnen bijdragen tot het verminderen van de invloed van stereotypen bij beslissingen.

Daarnaast richten de meeste studies over gestandaardiseerde toetsen zich op het meso- en macroniveau. Onderzoek naar hoe er met de resultaten van deze toetsen wordt omgegaan, blijft vaak beperkt tot feedbackrapporten op schoolniveau. Er is dus nog weinig gekend over hoe leerkrachten deze resultaten gebruiken op leerlingniveau. Ook binnen sociologisch onderzoek naar gelijke onderwijskansen ligt de focus sterk op het macroniveau. Voorbeelden hiervan zijn onderzoek naar vroege versus late tracking, schoolsegregatie etc. Ook hier gaat er minder aandacht naar de processen die zich afspelen in de klas. Het is daarom belangrijk om te verkennen op welke manier de beslissingsprocessen van leraren kunnen bijdragen aan het aanpakken van deze ongelijkheid.

1.3. Huidig onderzoek

Het doel van deze masterproef is om zicht te krijgen op hoe het gebruik van data uit gestandaardiseerde toetsen kan bijdragen tot gelijke onderwijskansen voor leerlingen met een lage SES. Dit onderzoek richt zich specifiek op hoe leerkrachten de data uit netgebonden toetsen gebruiken bij het al dan niet toekennen van een getuigschrift basisonderwijs, en de invloed hiervan op gelijke onderwijskansen.

In het huidige onderzoek worden de netgebonden toetsen gebruikt als voorbeeld van gestandaardiseerde toetsen in Vlaanderen. Het begrip 'gelijke onderwijskansen' wordt

geoperationaliseerd als de percepties tegenover GOK-leerlingen. Het huidige onderzoek focust zich op het beslissingsproces van het al dan niet toekennen van het getuigschrift basisonderwijs.

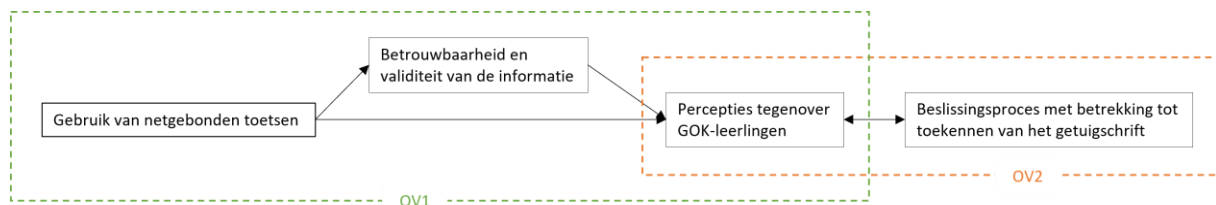
In de literatuur zijn er aanwijzingen te vinden dat gestandaardiseerde toetsen zorgen voor meer objectieve, betrouwbare informatie. Dit zou tegenwicht kunnen bieden aan de subjectieve stereotypen die leerkrachten soms hanteren. Daarom wordt dit concept als mediërende variabele opgenomen in het kader. Figuur 1 toont de samenhang tussen deze verschillende concepten.

Het huidige onderzoek zal een antwoord formuleren op volgende twee onderzoeksvragen:

- OV1.** In welke mate is er een verband tussen het gebruik van netgebonden toetsen en de percepties van leerkrachten tegenover GOK-leerlingen?
- OV2.** Op welke manier maken leerkrachten basisonderwijs gebruik van de resultaten van netgebonden toetsen op leerlingniveau bij de beslissing over het getuigschrift?

Figuur 1

Samenhang tussen de concepten



2. Theoretisch kader

In dit hoofdstuk worden de verschillende concepten uit de onderzoeksvragen in de diepte verkend. Eerst gaan we in op een kader voor beslissingsprocessen, waar het toekennen van het getuigschrift een voorbeeld van is. Vervolgens verdiepen we ons in de literatuur met betrekking tot de invloed van SES en GOK-indicatoren op het studieadvies van leerlingen. Daarna bekijken we wat de voor- en nadelen zijn van het resultaat van gestandaardiseerde toetsen in dit beslissingsproces te betrekken. Ten slotte gaan we dieper in op de netgebonden toetsen als operationalisering van gestandaardiseerde toetsen in Vlaanderen.

2.1. Beslissingsprocessen

2.1.1. Kader voor datagebaseerde beslissingen

Leerkrachten nemen veel beslissingen in de loop van de schoolloopbaan van een leerling. Dit doen ze op basis van data die ze via zowel rationele als intuïtieve processen verzamelen (Vanlommel, 2018). Bij rationele processen gaan leerkrachten deze data doelgericht en systematisch verzamelen, analyseren en interpreteren (Schildkamp & Ehren, 2013; Vanlommel et al., 2018). Daarnaast kunnen leerkrachten ook intuïtief patronen herkennen en interpreteren op basis van hun ervaring (Vanlommel, 2018). Dit gebeurt dan eerder spontaan en minder systematisch.

De literatuur onderscheidt een aantal fases die leraren doorlopen bij het nemen van een datagebaseerde beslissing (Coburn & Turner, 2011; Schildkamp et al., 2016; Vanlommel, 2018). De cyclus begint met het definiëren van een probleem waarmee aan de slag wordt gegaan. Vervolgens wordt er data verzameld, die in de fase daarna wordt geanalyseerd en geïnterpreteerd. Ten slotte worden de verschillende alternatieven naast elkaar gelegd en geëvalueerd. De rationele en intuïtieve processen lopen parallel doorheen de fases van het beslissingsproces (Vanlommel, 2018). Dit maakt dat rationeel verzamelde data op een intuïtieve manier geïnterpreteerd kunnen worden, of dat persoonlijke criteria een rol kunnen spelen in het belang dat aan bepaalde informatie wordt toegekend (Vanlommel, 2018; Vanlommel et al., 2021).

2.1.2. Getuigschrift basisonderwijs

Het beslissingsproces waar het huidige onderzoek op focust, is het toekennen van het getuigschrift basisonderwijs. Regelmatige leerlingen die de eindtermgerelateerde leerplandoelen in voldoende mate bereiken, krijgen aan het einde van het basisonderwijs een getuigschrift (Vlaams ministerie van onderwijs en vorming, 2016). Met 'in voldoende mate' wordt zowel het kwantitatieve aantal doelen als de kwalitatieve mate van beheersing bedoeld (Katholiek Onderwijs Vlaanderen, 2022a). Leerlingen die hun getuigschrift basisonderwijs behalen, kunnen in het secundair onderwijs starten in het eerste leerjaar A. Leerlingen die hun getuigschrift niet behalen, starten in het eerste leerjaar B (GO!, 2022).

De beslissing over het al dan niet toekennen van het getuigschrift wordt genomen door de klassenraad basisonderwijs (Vlaams ministerie van onderwijs en vorming, 2016). De klassenraad bestaat uit de personeelsleden die verantwoordelijk zijn voor het onderwijs aan een bepaalde groep leerlingen (Katholiek Onderwijs Vlaanderen, 2022a). Idealiter worden de leraren van de derde graad, de leraar lichamelijke opvoeding, de zorgcoördinator, leraren met een zorg- of GOK-opdracht en de directeur betrokken. Overall waar in het huidige onderzoek gesproken wordt van 'leraren', worden bij uitbreiding alle leden van de klassenraad basisonderwijs bedoeld.

Zowel het GO! (2022) als Katholiek Onderwijs Vlaanderen (2022a) geven aan dat de klassenraad rekening moet houden met het totaalbeeld van het kind, het groeipotentieel, de zelfstandigheid en de context waarin de leerling zich bevindt. Hierbij bekijkt de klassenraad niet alleen de resultaten op eindtoetsen, omdat deze resultaten slechts één informatiebron zijn naast vele andere. Het uitrekenen van het getuigschrift basisonderwijs mag niet louter gebaseerd zijn op de punten op leergebieden. Men moet ook rekening houden met het proces dat tot de resultaten heeft geleid, de persoonsontwikkeling en de mogelijkheden van het kind (GO!, 2022).

2.2. Percepties tegenover GOK-leerlingen

2.2.1. Invloed van SES op studieadvies

Het studieadvies dat leerlingen krijgen, wordt in zekere mate beïnvloed door hun SES. Leerlingen met een lagere SES krijgen, ongeacht eerdere schoolprestaties, minder vaak het advies om theoretisch georiënteerde studierichtingen te volgen dan leerlingen met een hogere SES (Boone & Van Houtte, 2013; Timmermans et al., 2015; van Leest et al., 2021). Hoe lager de SES, hoe kleiner de kans dat leerlingen een advies krijgen voor Latijn of moderne wetenschappen (Boone & Van Houtte, 2013). Experimenteel onderzoek toont aan dat wanneer twee leerlingen identieke prestaties voorleggen, de leerling met een hogere SES vaker wordt bevoordeeld ten opzichte van de leerling met een lagere SES (Geven et al., 2018). Ook op schoolniveau is er een effect van SES op oriëntatie. Ongeacht de eerdere prestaties van leerlingen, liggen de studieadviezen doorgaans hoger op scholen met een leerlingenpopulatie met een hogere SES (van Leest et al., 2021).

Glock et al. (2015) nuanceren bovenstaande bevindingen. Volgens dit onderzoek zijn de studieadviezen voor leerlingen uit etnische minderheidsgroepen vooral minder nauwkeurig dan bij meerderheidsgroepen. Leerlingen uit minderheidsgroepen krijgen vaker zowel hogere als lagere studieadviezen dan op basis van hun academisch profiel verwacht zou worden. In tegenstelling tot ander onderzoek, vonden van Leest et al. (2021) het effect van SES terug bij zowel beslissingen gebaseerd op het oordeel van leerkrachten als bij beslissingen gebaseerd op gestandaardiseerde toetsen. Voor beslissingen gebaseerd op het oordeel van leerkrachten was er echter een direct effect van SES op het advies dat niet aanwezig was bij beoordelingen gebaseerd op gestandaardiseerde toetsen.

Een groot deel van de leerkrachten geeft aan dat naast de schoolprestaties, talenten en interesses van leerlingen, ook het gedrag een belangrijk criterium is bij het formuleren van een studieadvies (Boone & Van Houtte, 2013). Leerkrachten kijken bijvoorbeeld naar de zelfredzaamheid, het planningsvermogen en de punctualiteit van de leerling. Informatie over deze vaardigheden wordt vaak op een eerder intuïtieve manier verzameld (Vanlommel et al., 2018). Volgens Dunkake & Schuchart (2015) benadeelt dit leerlingen met een lagere SES aangezien zij deze vaardigheden vaak minder van thuis meekrijgen, ongeacht hun cognitieve capaciteiten. Daarnaast worden leerlingen met een lagere SES volgens hetzelfde onderzoek gemiddeld als agressiever, minder ambitieus, luier en lastiger, ongedisciplineerd, slecht presterend en ongemotiveerd gepercipieerd. Leerkrachten die geloven in deze stereotypen verwachten vaak grotere verschillen in de prestaties van diverse leerlingengroepen binnen de klas. Lorenz (2021) verwacht dat stereotypen de beoordeling van leerkrachten over individuele leerlingen kunnen beïnvloeden.

Scholen houden bij het formuleren van een studieadvies ook vaak rekening met de verwachte steun en betrokkenheid van de ouders (Geven et al., 2018). Ouders met een hoge SES worden vaak in staat geacht om hun kinderen beter te helpen met schoolgerelateerde zaken, wat kan zorgen voor sociale verschillen. Daarnaast baseren leerkrachten hun oordeel over de ouderbetrokkenheid vaak op de betrokkenheid bij activiteiten op school, eerder dan op activiteiten thuis (Bakker et al., 2007).

Ook hier hebben gezinnen met een hogere SES vaker een streepje voor op gezinnen met een lagere SES. Ten slotte worden ouders vaak betrokken in het traject dat leidt tot het studieadvies. Onderzoek van Barg (2013) toont aan dat verschillen in studieadviezen van leerkrachten naar sociale achtergrond sterk verminderd worden wanneer rekening wordt gehouden met de voorkeur van de gezinnen zelf.

2.2.2. GOK-indicatoren

De Vlaamse overheid hanteert een aantal indicatoren bij het toekennen van extra middelen aan scholen in functie van een gelijke kansenbeleid. Op basis van het aantal indicatoren waar hun leerlingenpopulatie gemiddeld op aantikt, krijgt elke school een OKI-cijfer (Onderwijskansarmoede-indicator) tussen 0 en 4 (Katholiek Onderwijs Vlaanderen, 2021a). Leerlingen die voldoen aan één van deze criteria worden GOK-leerlingen genoemd. In het huidige onderzoek zullen dezelfde indicatoren gebruikt worden om de leerlingenkenmerken te omschrijven die een negatieve invloed hebben op de gelijke onderwijskansen van de leerling. De termen *OKI-indicatoren* en *GOK-leerlingen* behoren tot de gemeenschappelijke taal in het Vlaamse onderwijs wanneer we spreken over gelijke onderwijskansen. Daarnaast is er voor elk van deze indicatoren evidentie uit de literatuur wat betreft hun invloed op de gelijke onderwijskansen.

De eerste indicator is *thuis een andere taal spreken dan het Nederlands*. Thuis een andere taal spreken is volgens onderzoek van August et al. (2005) een belangrijke oorzaak van lagere scores op begrijpend lezen. Daarnaast scoren leerlingen met een andere thuistaal vaak ook lager op wetenschappen (Maerten-Rivera et al., 2010; Van Laere et al., 2014). De invloed van de thuistaal wordt volgens Agirdag en Korkmazer (2015) overschat. Anderstalige leerlingen kunnen hun achterstand inhalen (Broomes, 2013) en de ontwikkeling van de moedertaal kan het Nederlands ook net ten goede komen (Prevoe et al., 2015).

De tweede indicator is het *opleidingsniveau van de moeder*. Volgens Bourdieu (1970) is het onderwijs gemodelleerd naar de cultuur van de dominante sociale klasse, waardoor leerlingen uit lagere sociale klassen vaak niet beschikken over het culturele kapitaal dat nodig is om succesvol te zijn op school (Boone, 2013). De thuiscontext speelt een belangrijke rol bij de overdracht van cultureel kapitaal, waarbij ouders hun kinderen bewust of onbewust blootstellen aan bepaalde normen en waarden (Jæger, 2011). Kinderen uit hogere sociale klassen hebben vaak al een beter begrip van de verwachtingen van de school (Calarco, 2011), en ouders met een hogere SES zijn vaak beter in staat om te communiceren met leerkrachten en directies (Barg, 2012; Lareau, 2015).

De derde indicator is het *ontvangen van een schooltoelage*. Onderzoeken die armoede onderscheiden van SES laten zien dat opgroeien in armoede een negatief effect kan hebben op de schoolprestaties van leerlingen. Het beïnvloedt zowel het fysieke als mentale welzijn en kan leiden tot een verminderde schoolrijpheid (Ferguson et al., 2007). Michelmores en Dynarski (2017) ontdekten een negatief, bijna lineair verband tussen het aantal jaren dat een leerling in economisch achtergestelde omstandigheden leefde en de testresultaten.

De vierde en laatste indicator is *wonen in een buurt met veel schoolse vertraging*. Buurtarmoede heeft volgens een reviewstudie van Nieuwenhuis en Hooimeijer (2015) een negatief effect op de onderwijsprestaties van leerlingen, zelfs na controle voor andere factoren zoals de SES van individuele families. Bsmetting, waarbij bewoners worden beïnvloed door het gedrag en de houding van hun buren, wordt genoemd als een verklarend sociaal mechanisme voor buurteffecten. Daarnaast speelt de sociaal kapitaaltheorie een rol, waarbij gezinnen met een hoge SES doorgaans toegang hebben tot meer ondersteuning via hun netwerken, terwijl gezinnen in buurten met veel schoolse vertraging vaak minder steun kunnen vinden in hun netwerk (Boone, 2013).

2.3. Betrouwbaarheid en validiteit van de beschikbare informatie

2.3.1. Valkuilen bij evaluaties door leerkrachten

De beoordelingen die leerkrachten aan leerlingen geven, voldoen vaak niet aan vereisten op vlak van betrouwbaarheid, validiteit en objectiviteit. Zo kennen verschillende leerkrachten bijvoorbeeld een verschillende waarde toe aan een bepaald cijfer, wat vergelijken moeilijk maakt (Allal, 2013). In hun reviewstudie beschrijven Geven et al. (2018) dat leerkrachten geneigd zijn systematische fouten te maken in hun beoordelingen van specifieke groepen leerlingen. Naast de prestaties van de leerlingen zijn er ook demografische kenmerken die de beoordeling van de leerkracht beïnvloeden.

Hoewel prestaties een belangrijke factor zijn bij deze beoordeling, is dit niet het enige waar leerkrachten zich op baseren. Uit onderzoek naar de beoordeling van leerlingprestaties door leerkrachten blijkt dat zij zowel prestatiegebonden als niet-prestatiegebonden factoren door elkaar gebruiken (Brookhart, 2013). Dit maakt dat het oordeel dat zij vellen over de prestaties van leerlingen vaak minder valide is.

Wanneer leerkrachten het gevoel hebben dat ze onvoldoende informatie hebben om de capaciteiten van een leerling te beoordelen, zijn ze sneller geneigd om groepskenmerken te gebruiken om het potentieel van de leerling in te schatten. Hierbij worden de groepskenmerken gebruikt als proxy om het beeld van de leerkracht te vervolledigen (Geven et al., 2018). Daarnaast kan het ook gebeuren dat leerkrachten sommige leerlingen anders behandelen op basis van de groep waartoe zij behoren. Bepaalde vooroordelen ten opzichte van deze specifieke groepen liggen aan de basis hiervan. Deze vooroordelen kunnen zowel impliciet als expliciet zijn. Leerkrachten zijn zich dus niet altijd bewust van de stereotypen van waaruit ze handelen.

2.3.2. Kansen bij het gebruik van gestandaardiseerde toetsen

Om bovengenoemde redenen kan er volgens Brookhart (2013) voor high-stakes, summatieve beoordelingen beter op gestandaardiseerde toetsen vertrouwd worden dan op het oordeel van de leerkracht. Een gestandaardiseerde toets houdt in dat leerlingen een toets afleggen over dezelfde inhoud, in identieke omstandigheden (Penninckx et al., 2017). Dit maakt het mogelijk om scores tussen leerlingen en scholen te vergelijken.

Door gebruik te maken van de informatie uit gestandaardiseerde toetsen kunnen leerkrachten hun studieadviezen meer baseren op objectieve standaarden en minder op subjectieve zaken zoals impliciete stereotypen en de voorkeuren van ouders (van de Werfhorst, 2015). Op deze manier zullen leerkrachten hun studieadvies meer baseren op de feitelijke prestaties van de leerling (Bol et al., 2014; Wyatt-Smith & Klenowski, 2013). Objectievere informatie kan de onzekerheid van een leerkracht over het onderwijspotentieel van een kind verminderen (Reskin, 2003).

Gestandaardiseerde toetsen kunnen er ook voor zorgen dat de verantwoording voor scholen (accountability) toeneemt. Mensen zijn gemotiveerder om een correcte beslissing te nemen wanneer zij voor deze beslissing verantwoordelijk worden gehouden (Koch et al., 2015). Het verhogen van het verantwoordelijkheidsgevoel van leerkrachten voor hun adviezen kan de invloed van leerlingkenmerken verkleinen (Koch et al., 2015; Pit-ten Cate et al., 2016). Wanneer leerkrachten verplicht worden om hun adviezen te baseren op scores van gestandaardiseerde toetsen, zullen zij minder snel geneigd zijn om te discrimineren (Reskin, 2003).

2.3.3. Valkuilen bij het gebruik van gestandaardiseerde toetsen

Toch is ook data uit gestandaardiseerde toetsen niet per definitie betrouwbaar en valide. Uit onderzoek van Lopez-Agudo et al. (2021) blijkt bijvoorbeeld dat meertalige leerlingen benadeeld worden bij het afnemen van de PISA-toetsen. Leerlingen die de toets afnemen in een andere taal dan hun moedertaal scoren beduidend lager dan leerlingen die de toets wel in hun moedertaal afnemen. Doordat zij de inhoud van de toets niet altijd volledig begrijpen, meet de toets hun eigenlijke competenties niet op een nauwkeurige manier.

Testscores op gestandaardiseerde toetsen kunnen in sommige gevallen meer beïnvloed worden door context- en achtergrondkenmerken van leerlingen dan door individuele capaciteiten (Au, 2015). Gestandaardiseerde toetsen zouden de onderwijsongelijkheid op die manier eerder versterken dan doen afnemen. Daarnaast beschikken leerlingen uit deze groepen vaak over een lagere zelf-effectiviteit, wat hun toetsscores negatief kan beïnvloeden (Rogers-Ard et al., 2013). Daarom zijn gestandaardiseerde toetsen volgens Au (2015) minder geschikt om te gebruiken voor high-stakes beslissingen.

Enkel gebruik maken van gestandaardiseerde toetsen voor de beoordeling van leerlingen geeft de leerkrachten daarnaast een onvolledig beeld over de capaciteiten van een leerling. De interacties met de leerlingen in de dagelijkse klaspraktijk kunnen leerkrachten een rijker zicht geven op de capaciteiten van een leerling dan toetsen kunnen meten (Brookhart, 2013). Leraren zoeken vaker alternatieve, kwalitatieve informatiebronnen als de resultaten van een leerling op gestandaardiseerde toetsen fluctueren of als ze niet overtuigd zijn dat deze een correct beeld geven van de capaciteiten van de leerling (Allal, 2013). Hoewel de leerkrachten zorgen voor enige standaardisatie van hun beoordelingsprocedures, doen ze ook aanpassingen om rekening te houden met de individuele kenmerken van elke leerling.

2.3.4. Complementaire vormen van beoordelen

Het is voor een leraar belangrijk om alle beschikbare informatie over een leerling te gebruiken om het meest geschikte oordeel te vellen voor die specifieke leerling. Allal (2013) pleit er daarom voor gestandaardiseerde toetsen en beoordelingen door leraren te beschouwen als twee verschillende maar complementaire vormen van beoordeling. Ze zijn elk op zich legitiem en moeten elk voldoen aan hoge, maar niet-identieke, kwaliteitsnormen. De intuïtie van leerkrachten kan hen sturen om te bepalen welke informatie ze al dan niet gebruiken, en om de informatie zoals toetsresultaten te interpreteren (Vanlommel et al., 2018).

2.4. Gebruik van netgebonden toetsen

Scholen zijn verplicht om hun leerlingen aan het einde van de lagere school te laten deelnemen aan gevalideerde toetsen (Vlaams ministerie van onderwijs en vorming, 2017). Deze toetsen moeten minstens drie leergebieden omvatten. Scholen kunnen vrij kiezen tussen de netgebonden toetsen van Katholiek Onderwijs Vlaanderen (interdiocesane proeven) of Onderwijskoepel van Steden en Gemeenten (OVSG-toetsen), of voor de paralleltoetsen van de peilingsproeven die ontwikkeld zijn in opdracht van het Vlaams Ministerie van Onderwijs en Vorming. In het huidige onderzoek baseren we ons op de netgebonden toetsen.

2.4.1. Situering van de netgebonden toetsen

Zowel Katholiek Onderwijs Vlaanderen als Onderwijskoepel van Steden en Gemeenten (OVSG) ontwikkelden gevalideerde toetsen, respectievelijk de interdiocesane proeven (IDP) en de OVSG-toets. Scholen kunnen in theorie vrij kiezen met welke toets ze aan de slag gaan om aan hun decretale verplichting te voldoen, maar zijn in de praktijk erg trouw aan de eigen koepel. In juni 2016 nam 85% van de katholieke scholen deel aan IDP; 98% van OVSG-scholen en 92% van de GO!-scholen nam deel aan de OVSG-toets (Steunpunt Toetsontwikkeling en Peilingsonderzoek [STEP], 2017).

Beide toetsen baseren zich op de eigen leerplannen, namelijk ZILL voor de IDP en de leerplannen van OVSG voor de OVSG-toets. Katholiek Onderwijs Vlaanderen (2022b) geeft aan een breed aanbod te voorzien met toetsen voor de leergebieden mens & maatschappij, Nederlands, wetenschappen & techniek, wiskunde, Frans en muzische vorming. Later zal hier ook lichamelijke opvoeding aan worden toegevoegd. Ze kiezen voor een scherpe focus met een beperkt aantal leerinhouden per leerdoel. Hierdoor beogen ze de inhouden meer in de diepte te evalueren en zo zinnigere informatie te verkrijgen. OVSG (2023) noemt haar gestandaardiseerde toets een “alle-talenten-toets” die inzet op een brede evaluatie. De toets meet vanuit het leerplan alle domeinen van de eindtermen. Hierbij is er aandacht voor zowel kennis als vaardigheden (OVSG, 2015; OVSG, 2023).

2.4.2. Gebruik op school- en leerlingniveau

Voor beide toetsen ligt de focus voornamelijk op het versterken van de interne kwaliteitszorg van de scholen (Penninckx et al., 2017). Beide koepels geven aan dat de toetsen de deelnemende scholen kunnen helpen om een zicht te krijgen op de sterktes en groeikansen binnen de schoolwerking (Katholiek Onderwijs Vlaanderen, 2021c; OVSG, 2015). Scholen ontvangen hiervoor bij zowel Katholiek Onderwijs Vlaanderen als OVSG een rapport op school- en klasniveau. Scholen kunnen via dit rapport de eigen scores vergelijken met het Vlaamse gemiddelde en het gemiddelde van de referentiegroep. Deze referentiegroep bestaat uit scholen met een vergelijkbare populatie volgens de onderwijs kansarmoede-indicator (Katholiek Onderwijs Vlaanderen, 2022b; OVSG, 2023; STEP, 2017).

Naast het school- en klasniveau kan de OVSG-toets ook op leerlingniveau gebruikt worden. De toets bevraagt de leerplandoelen en kan dienst doen als informatiebron bij de evaluatie van leerlingen. OVSG geeft echter aan dat de toets niet op zichzelf gebruikt mag worden voor het toekennen van een getuigschrift, maar altijd naast schooleigen gegevens gelegd moet worden (OVSG, 2015). Katholiek Onderwijs Vlaanderen benadrukt vooral dat de toetsen niet ontwikkeld zijn om een eindoordeel te vellen over het al dan niet bereiken van de doelstellingen. Daarnaast zien ze wel mogelijkheden om de resultaten van de toets aanvullend op andere evaluaties te gebruiken (Katholiek Onderwijs Vlaanderen, 2021b). Dit verschil in visie en communicatie is ook merkbaar in het gebruik van de toetsen door scholen. Onderzoek van het Steunpunt Toetsontwikkeling en Peilingsonderzoek (2017) toont aan dat directeurs van OVSG-scholen vaker gebruik op leerlingniveau rapporteren dan directeurs van Katholieke scholen.

3. Methodologie

3.1. Onderzoeksdesign

De twee onderzoeksvragen van het huidige onderzoek worden beantwoord aan de hand van een concurrent mixed method design waarbij de kwantitatieve en kwalitatieve data gelijktijdig werden verzameld (Almeida, 2018). Dit heeft als voordeel dat de twee dataverzamelingen elkaar kunnen aanvullen en elkaars zwakke punten kunnen compenseren. Het kwantitatieve luik geeft inzicht in algemene verbanden en het kwalitatieve luik kan de processen achter deze verbanden meer in de diepte beschrijven.

De eerste onderzoeksvraag wordt beantwoord aan de hand van een surveyonderzoek, de tweede onderzoeksvraag aan de hand van semigestructureerde interviews. De resultaten van beide dataverzamelingen worden apart gerapporteerd in het resultatenluik en vervolgens geïntegreerd in de discussie en conclusie.

3.2. Kwantitatief luik (OV1)

3.2.1. Respondenten

De populatie van het onderzoek bestaat uit de leden van de klassenraad basisonderwijs die betrokken zijn bij de beslissing over het al dan niet toekennen van het getuigschrift. Eind februari 2023 werden directeurs van 100 willekeurige scholen aangeschreven met de vraag om de oproep tot deelname te verspreiden binnen hun school. Gezien de focus van dit onderzoek werden enkel basis- en lagere scholen gecontacteerd. Secundaire scholen en autonome kleuterscholen werden niet weerhouden.

Om een voldoende grote steekproef trachten te bekomen, werd er begin maart 2023 ook een oproep tot deelname gedaan via sociale media. De oproep werd gedeeld op zowel Facebook, Twitter als LinkedIn om een zo breed en verscheiden mogelijke groep respondenten aan te spreken.

De survey werd door 99 respondenten ingevuld. Eén respondent met onvolledige data werd verwijderd, wat de uiteindelijke steekproef op 98 respondenten bracht. De steekproef bestond voor het grootste deel ($n = 72$) uit vrouwen. Het gesubsidieerd vrij onderwijs was met het meest vertegenwoordigd ($n = 66$), gevolgd door het gesubsidieerd officieel onderwijs ($n = 23$) en het gemeenschapsonderwijs ($n = 9$). De grootste groep respondenten was lid van het directieteam ($n = 32$) of van het zorgteam ($n = 31$). Daarnaast namen ook een aantal klasleerkrachten 6^e leerjaar ($n = 13$) en andere leerkrachten ($n = 20$) deel aan de survey. Tenslotte gaven twee respondenten aan een andere functie op te nemen. De respondenten voerden hun huidige functie op het moment van deelname tussen de 1 en 38 jaar uit, met een mediaan van 6 jaar.

3.2.2. Dataverzameling

De gebruikte survey bestond uit drie schalen van gevalideerde vragenlijsten, aangevuld met één vraag die peilde naar de mate waarin de respondent gebruik maakt van netgebonden toetsen bij het al dan niet toekennen van een getuigschrift basisonderwijs. Op die manier volgde de survey dezelfde structuur als het conceptueel kader dat wordt afgebeeld in Figuur 1. De opbouw van de survey wordt weergegeven in Tabel 1. De volledige survey is bijgevoegd in Bijlage 2.

Tabel 1*Opbouw van de survey*

Onderdeel	Items	Cronbach's α	Voorbeeldvraag
Mate van gebruik netgebonden toetsen	1	nvt	In welke mate maakt u gebruik van gestandaardiseerde toetsen, ontwikkeld binnen een koepel (interdiocesane proeven, OVSG-toets...) bij het nemen van beslissingen over het al dan niet toekennen van een getuigschrift basisonderwijs aan een leerling?
Betrouwbaarheid en validiteit van de informatie	5	0.84	De informatie die ik ter beschikking heb, is objectief.
Negatieve percepties tegenover GOK-leerlingen	5	0.86	GOK-leerlingen spannen zich gemiddeld minder in op school dan andere leerlingen.
Sociaal wenselijkheid	5	0.57	Met wie ik ook praat, ik kan altijd goed luisteren.

De items van de schalen *Betrouwbaarheid en validiteit van de informatie* en *Negatieve percepties tegenover GOK-leerlingen* werden beantwoord aan de hand van een Likertschaal met vijf antwoordmogelijkheden, namelijk *Helemaal oneens*, *Eerder oneens*, *Noch oneens/noch eens*, *Eerder eens*, en *Helemaal eens*. Er werd daarnaast ook steeds een optie *Weet niet/niet van toepassing* voorzien. Voor het item over de mate van gebruik van de netgebonden toetsen waren er vijf categorieën als antwoordmogelijkheden.

Voor de schaal *Betrouwbaarheid en validiteit van de informatie* werd gebruik gemaakt van een schaal uit een onderzoek naar informatiegebruik door schoolleiders en leerkrachten (Van Gasse et al., 2015). Deze schaal werd integraal overgenomen. De Cronbach's alpha van 0.84 in het huidige onderzoek wijst op een goede interne consistentie (George & Mallery, 2003, geciteerd in De Maeyer et al., 2020).

Voor de schaal in verband met *Negatieve percepties tegenover GOK-leerlingen* werd een schaal uit de vragenlijst van Lorenz (2021) herwerkt. De schaal werd uit het Engels vertaald naar het Nederlands en de doelgroep 'leerlingen met een migratieachtergrond' werd vervangen door 'GOK-leerlingen'. De Cronbach's alpha van 0.86 in het huidige onderzoek wijst ook voor deze schaal op een goede interne consistentie (George & Mallery, 2003, geciteerd in De Maeyer et al., 2020).

De schalen uit de twee vragenlijsten zijn reeds getest en hebben bewezen betrouwbaar en valide te zijn. De aanpassingen aan de vragenlijst van Lorenz (2021) kunnen de betrouwbaarheid en validiteit beïnvloeden ten opzichte van het oorspronkelijke onderzoek. De aanpassingen verhogen echter de inhoudsvaliditeit door de terminologie in de vragen meer af te stemmen op de onderzochte constructen. Ook in het huidige onderzoek is de interne consistentie van deze schalen goed.

De cognitieve validiteit werd nagegaan aan de hand van een verbaal report (Leighton, 2004). Er werd aan één proefrespondent gevraagd om de vragenlijst in te vullen en hierbij alle gedachten luidop te verwoorden. Dit gaf de onderzoeker gedeeltelijk inzage in het denkproces van de respondent, bijvoorbeeld op vlak van interpretatie van de vragen. Omwille van de haalbaarheid werd deze proefafname bewust zeer klein gehouden. Er werden op basis van de proefafname geen wijzigingen gedaan aan de survey.

Aangezien de schaal over de *Negatieve percepties tegenover GOK-leerlingen* peilt naar attitudes, bestond het risico dat respondenten sociaal wenselijk zouden antwoorden. Het risico op sociaal

wenselijk antwoordgedrag werd gemeten aan de hand van een eigen vertaling van de SDRS-5 (Hays et al., 1989). Dit is een korte schaal van vijf items die lijken te peilen naar relaties met anderen. De schaal maakt gebruik van een vijf-puntschaal waarbij telkens één van beide extreme antwoordmogelijkheden als sociaal wenselijk wordt gezien. Omwille van de zwakke interne consistentie van de schaal in het huidige onderzoek ($\alpha = 0.57$) zijn de resultaten in het huidige onderzoek niet gebruikt om respondenten uit de dataset te verwijderen.

De survey liep gedurende maart 2023 en had een gemiddelde invultijd van 4 minuten en 15 seconden.

3.2.3. Analyse

Voor de schalen *Betrouwbaarheid en validiteit van de informatie* en *Negatieve percepties tegenover GOK-leerlingen* werd een schaalscore tussen 0 en 5 bekomen door het gemiddelde van de onderliggende items te berekenen. Voor de schaal *Betrouwbaarheid en validiteit van de informatie* betekende 0 een klein en 5 een groot gevoel van betrouwbaarheid en validiteit. Voor de schaal *Negatieve percepties tegenover GOK-leerlingen* betekende 0 weinig en 5 veel vooroordelen tegenover GOK-leerlingen.

De drie schalen werden eerst afzonderlijk geanalyseerd aan de hand van beschrijvende statistiek, zoals de parameters van ligging en spreiding die overeen komen met het meetniveau van de schaal. Vervolgens werden de verbanden tussen de verschillende variabelen in kaart gebracht.

De analyse van de verbanden tussen de concepten *Mate van gebruik van netgebonden toetsen*, *Betrouwbaarheid en validiteit van de informatie*, en *Negatieve percepties tegenover GOK-leerlingen* volgde de structuur van het conceptuele kader (zie Figuur 1). Er werd onderzocht wat het directe effect is van het gebruik van netgebonden toetsen op de negatieve percepties tegenover GOK-leerlingen, en welk deel van het totale effect kan worden verklaard door de gepercipieerde betrouwbaarheid en validiteit van de informatie als mediator.

Structural equation modeling (SEM) wordt doorgaans gezien als ideale analysetechniek voor complexere modellen zoals deze met een mediator (Gunzler et al., 2013). Modellen met een mediator vereisen echter grote steekproeven tot 450 respondenten wegens mogelijke vertekeningen van de parameterschattingen (Wolf et al., 2013). Gezien de kleinere steekproef in het huidige onderzoek werden de drie verbanden afzonderlijk van elkaar geanalyseerd. Aangezien zowel de variabele *Betrouwbaarheid en validiteit van de informatie* als *Negatieve percepties tegenover GOK-leerlingen* niet normaal verdeeld waren, werden alle drie de analyses gedaan aan de hand van een ranked based regression. Dit is een niet-parametrisch alternatief voor een regressieanalyse.

3.3. Kwalitatief luik (OV2)

3.3.1. Respondenten

De respondenten die de survey invulden, konden aan het einde aangeven of ze geïnteresseerd waren om deel te nemen aan een interview. Op deze vraag reageerden dertien personen positief. Ze lieten hun contactgegevens achter en ontvingen nadien een mail met meer informatie. Van deze groep respondenten zegden er uiteindelijk vijf toe voor het interview. Tabel 2 toont enkele relevante kenmerken van de respondenten. Alle respondenten waren vrouwen en gebruikten de interdiocesane proeven als netgebonden toets.

Tabel 2*Kenmerken van de geïnterviewde respondenten*

Respondent	Huidige functie	GOK-leerlingen ¹	Niet-GOK-leerlingen ¹
Respondent 1	Zorgcoördinator	1	2
Respondent 2	Zorgleerkracht en beleidsondersteuner	2	0
Respondent 3	Directie	0	1
Respondent 4	Zorgcoördinator	2	0
Respondent 5	Leerkracht L6	2	0

Noot. ¹ aantal leerlingen besproken tijdens het interview

3.3.2. Dataverzameling

De data voor het kwalitatieve onderzoeksluik werd verzameld aan de hand van semigestructureerde interviews. Een semigestructureerd interview combineert vooraf opgestelde vragen in een interviewleidraad met ad hoc vragen tijdens het interview zelf. Dit heeft als voordeel dat je als onderzoeker kan bewaken dat bepaalde vragen zeker aan bod komen, maar dat je tegelijk ook voldoende flexibiliteit hebt om bijvragen te stellen waar nodig (Wilson, 2014).

Omdat er gewerkt werd met een interviewleidraad, kwamen de belangrijke vragen in elk interview aan bod en werden ze telkens op eenzelfde manier gesteld. Dit beperkt de invloed van de interviewer in vergelijking met ongestructureerde interviews en verhoogt op die manier de betrouwbaarheid van de resultaten.

De vragen uit de interviewleidraad zijn gebaseerd op de thema's uit de literatuurstudie. Zo werden er open vragen gesteld die peilen naar onderwerpen zoals de informatie waarop de klassenraad zich baseert voor hun beslissing, ouderbetrokkenheid, leerlingenkenmerken etc. Er werd vooraf aan de respondenten gevraagd om na te denken over twee casussen waarbij er vorig jaar getwijfeld werd over het al dan niet toekennen van het getuigschrift. Op die manier kon er meer in de diepte gepeild worden naar specifieke ervaringen, in plaats van enkel naar algemene percepties. Gezien de insteek van het huidige onderzoek werd aan de respondenten gevraagd om, indien mogelijk, één casus met een GOK-leerling en één met een niet-GOK-leerling te kiezen. De volledige interviewleidraad is bijgevoegd in Bijlage 3.

De samenhang tussen de literatuurstudie en de interviewvragen verhoogt de validiteit van het onderzoek. Er werd één proefinterview afgenomen om na te gaan of de formulering van de vragen duidelijk was. Op basis hiervan werden enkele kleine aanpassingen gedaan aan de interviewleidraad. Ook de manier van werken aan de hand van concrete cases maakt dat er dieper kon worden ingegaan op wat het onderzoeksluik precies beoogt te meten.

De interviews gingen allemaal online door. Dit verhoogde de haalbaarheid voor zowel de onderzoeker als de respondent, en maakte dat de respondenten meer op hun gemak waren dan wanneer ze in een vreemde setting geïnterviewd zouden worden. De interviews vonden plaats eind maart 2023 en duurden tussen de 36 en 51 minuten, met een gemiddelde van 45 minuten.

3.3.3. Analyse

De interviews werden met toestemming van de respondenten opgenomen en nadien ‘intelligent verbatim’ getranscribeerd. Dit is een vorm van transcriberen waarbij het transcript beter leesbaar wordt gemaakt door dialecten, herhalingen etc. niet letterlijk over te nemen (McMullin, 2023). Non-verbale signalen werden wel mee opgenomen indien relevant. In Bijlage 4 werd een voorbeeld van een transcript bijgevoegd.

De transcripten werden vervolgens gecodeerd in NVivo. Er werd zowel deductief als inductief gecodeerd. Het deductief coderen gebeurde aan de hand van een vooraf opgestelde codeboom, gebaseerd op de onderwerpen uit de literatuurstudie. Deze codeboom werd nadien inductief aangevuld met nieuwe informatie of thema's uit de interviews. Deze codes hadden betrekking tot thema's zoals de achtergrondkenmerken van de leerling, op welke manier de verschillende informatiebronnen gebruikt werden en welke informatie meespeelde in de eindbeslissing. De volledige codeboom is bijgevoegd in Bijlage 5.

4. Onderzoeksresultaten

4.1. Onderzoeksvraag 1

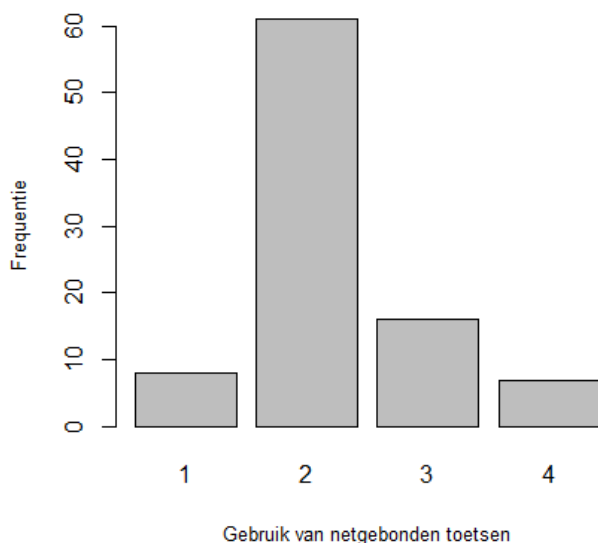
De eerste onderzoeksvraag naar het verband tussen het gebruik van netgebonden toetsen en de percepties van leerkrachten tegenover GOK-leerlingen, werd gemeten aan de hand van de kwantitatieve dataverzameling. De rapportering van dit onderzoeksluik begint met beschrijvende statistiek van de verschillende variabelen. Voor elke variabele werden de parameters van ligging en spreiding geanalyseerd. Vervolgens wordt er dieper ingegaan op de samenhang tussen deze variabelen.

4.1.1. Beschrijvende statistiek

De meerderheid van de respondenten gaf aan vooral gebruik te maken van de schooleigen toetsen voor de beslissing over het al dan niet toekennen van het getuigschrift. De gestandaardiseerde toetsen worden hierbij vooral aanvullend gebruikt ($Mdn = 2$, $IQR = 0$). Geen enkele respondent duidde aan enkel de netgebonden toetsen te gebruiken. Figuur 2 toont de verdeling van de respondenten over de verschillende antwoordcategorieën.

Figuur 2

Gebruik van eigen en netgebonden toetsen

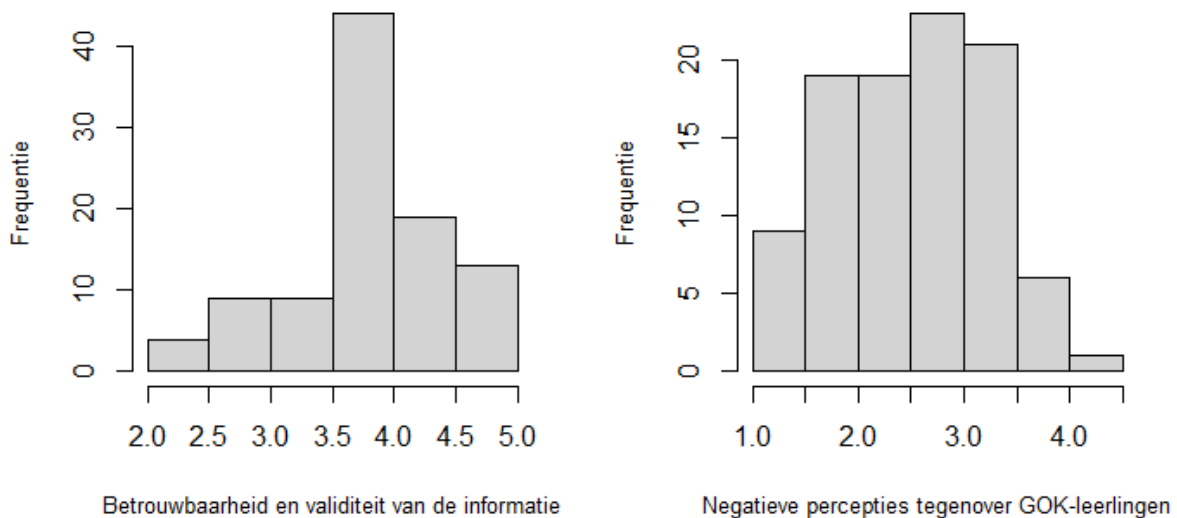


Noot. Categorieën: 1 = Enkel eigen toetsen; 2 = Vooral eigen toetsen, aangevuld met netgebonden toetsen; 3 = Eigen en netgebonden toetsen evenveel; 4 = Vooral netgebonden toetsen, aangevuld met eigen toetsen

De variabele *Betrouwbaarheid en validiteit van de informatie* heeft een gemiddelde score van 3.86 ($SD = 0.63$). De variabele *Negatieve percepties tegenover GOK-leerlingen* heeft een gemiddelde score van 2.55 ($SD = 0.74$). Figuur 3 toont de frequentieverdeling van beide variabelen.

Figuur 3

Histogram van de schalen 'Betrouwbaarheid en validiteit van de informatie' en 'Negatieve percepties tegenover GOK-leerlingen'



4.1.2. Verbanden tussen de variabelen

Het verband tussen het gebruik van netgebonden toetsen, de betrouwbaarheid en validiteit van de informatie, en percepties tegenover GOK-leerlingen werd gemeten in drie stappen. Eerst werd het directe verband nagegaan tussen het *Gebruik van de netgebonden toetsen* en de *Negatieve percepties tegenover GOK-leerlingen*. Vervolgens werd het verband tussen het *Gebruik van netgebonden toetsen* en de *Betrouwbaarheid en validiteit van de informatie* nagegaan, en ten slotte het verband tussen de *Betrouwbaarheid en validiteit van de informatie* en de *Negatieve percepties tegenover GOK-leerlingen*.

Effect van de mate van het gebruik van de netgebonden toetsen op de negatieve percepties tegenover GOK-leerlingen

In de steekproef waren er enkele verschillen in de negatieve percepties tegenover GOK-leerlingen naargelang de mate waarin de leerkrachten gebruik maken van de netgebonden toetsen. Echter, zowel het intercept ($p = .33$) als de effecten van vooral de eigen toetsen ($p = .27$), beide evenveel ($p = .63$) als vooral de netgebonden toetsen ($p = .99$) te gebruiken, waren niet significant. Tabel 3 toont een overzicht van de regressiecoëfficiënten.

We verwachten dat de mate van het gebruik van de netgebonden toetsen in de populatie geen effect heeft op de negatieve percepties tegenover GOK-leerlingen. Het al dan niet gebruik maken van de netgebonden toetsen zal dus geen verschil maken in hoe de leraar kijkt naar GOK-leerlingen.

Tabel 3

Regressiecoëfficiënten voor het effect van de mate van het gebruik van de netgebonden toetsen op de negatieve percepties tegenover GOK-leerlingen

Variabele	Estimate	SE	t	p
Intercept	-0.47	0.48	-0.99	.33
Gebruik van toetsen				
Vooraf eigen toetsen	0.54	0.49	1.11	.27
Beide evenveel	0.27	0.57	0.47	.63
Vooraf netgebonden toetsen	0.01	0.71	0.01	.99

Noot. $R^2 = .02$, $p = .52$

Effect van de mate van het gebruik van de netgebonden toetsen op het gevoel van betrouwbaarheid en validiteit van de informatie

Respondenten die aangaven enkel eigen toetsen te gebruiken, vonden hun informatie gemiddeld betrouwbaar en valide. Zowel vooral de eigen toetsen gebruiken als beide toetsen evenveel gebruiken, had geen effect op het gevoel van betrouwbaarheid en validiteit van de informatie. Bij respondenten die aangaven vooral de netgebonden toetsen te gebruiken, daalde het gevoel van betrouwbaarheid en validiteit significant met 1.28 standaardafwijkingen ($p = .001$). De mate van gebruik van de netgebonden toetsen verklaarde een groot deel van de variantie in het gevoel over betrouwbare en valide informatie te beschikken ($R^2 = .15$). Dit effect was significant ($p = .002$). Tabel 4 toont een overzicht van de regressiecoëfficiënten.

We verwachten dat de mate van gebruik van de netgebonden toetsen in de populatie een groot effect heeft op het gevoel van betrouwbaarheid en validiteit van de informatie. Specifiek het gebruik van vooral de netgebonden toetsen, aangevuld met de eigen toetsen, heeft een negatief effect. Dit betekent dat leraren die vooral de netgebonden toetsen gebruiken, de betrouwbaarheid en validiteit van hun beschikbare informatie doorgaans lager inschatten dan anderen.

Tabel 4

Regressiecoëfficiënten voor het effect van de mate van het gebruik van de netgebonden toetsen op het gevoel van betrouwbaarheid en validiteit van de informatie

Variabele	Estimate	SE	t	p
Intercept	0.22	0.23	0.98	.33
Gebruik van toetsen				
Vooraf eigen toetsen	-0.00	0.26	0.00	1
Beide evenveel	-0.00	0.31	0.00	1
Vooraf netgebonden toetsen	-1.28	0.38	-3.37	.001

Noot. $R^2 = .15$, $p = .002$

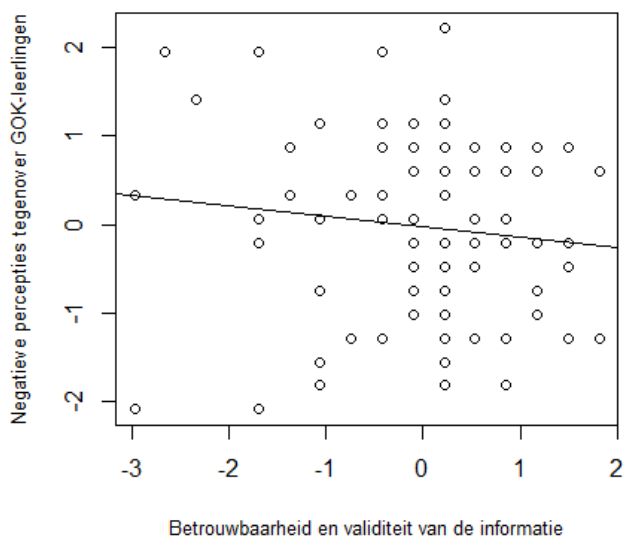
Effect van het gevoel van betrouwbaarheid en validiteit van de informatie op de negatieve percepties tegenover GOK-leerlingen

Bij een gemiddelde betrouwbaarheid en validiteit van de informatie waren de negatieve percepties tegenover GOK-leerlingen eveneens gemiddeld. In de steekproef zorgde een stijging op vlak van betrouwbaarheid en validiteit van de informatie voor een lichte daling in de negatieve percepties tegenover GOK-leerlingen. Dit effect was echter niet significant ($p = .26$). Het gevoel over betrouwbare en valide informatie te beschikken, verklaarde een klein deel van de variantie in negatieve percepties tegenover GOK-leerlingen ($R^2 = .01$). Dit effect was niet significant ($p = .43$). Figuur 4 toont de samenhang tussen de betrouwbaarheid en validiteit van de informatie en de negatieve percepties tegenover GOK-leerlingen.

We verwachten dat het gevoel van betrouwbaarheid en validiteit van de informatie in de populatie geen effect heeft op de negatieve percepties tegenover GOK-leerlingen. Dit betekent dat hoe betrouwbaar en valide leraren hun beschikbare informatie al dan niet vinden, geen invloed heeft op hoe zij naar GOK-leerlingen kijken.

Figuur 4

Samenhang tussen de betrouwbaarheid en validiteit van de informatie en de negatieve percepties tegenover GOK-leerlingen



Noot. Variabelen uitgedrukt in Z-scores

4.2. Onderzoeksvraag 2

De tweede onderzoeksvraag: “Op welke manier maakt de klassenraad basisonderwijs gebruik van de resultaten van netgebonden toetsen op leerlingniveau bij de beslissing over het getuigschrift?” wordt beantwoord aan de hand van de kwalitatieve data uit de semigestructureerde interviews. We gaan hierbij in op het beslissingsproces in het algemeen en benoemen daarnaast ook specifieke uitspraken die gelinkt zijn aan het onderwerp van gelijke onderwijskansen.

4.2.1. Achtergrondkenmerken van de leerlingen

De meeste voorbeelden van de respondenten gingen over GOK-leerlingen. Een vaak genoemde reden hiervoor was dat de respondenten in een school werken met een groot aantal GOK-leerlingen in het algemeen. De meerderheid van de respondenten gaf aan dat de leerlingen thuis een andere taal spraken dan het Nederlands. Volgens een aantal respondenten had dit een invloed op de latere leerproblemen en de twijfels met betrekking tot het getuigschrift. Eén respondent vertelde bijvoorbeeld over een anderstalige leerling die bij hen toekwam in het derde leerjaar en de taal nog moest verwerven.

Naast de thuistaal noemden enkele respondenten ook andere GOK-indicatoren zoals armoede of het diploma van de moeder. Bij een groot deel van de GOK-leerlingen gaven de respondenten meerdere indicatoren aan. Aanvullend op deze indicatoren gaven sommige respondenten nog verdere informatie over de thuissituatie van deze leerlingen. Zo moest één leerling bijvoorbeeld thuis voor de jongere broers en zussen zorgen, en was er in een ander gezin sprake van druggebruik en fysiek geweld.

Ja ja kansarmoede, dat geeft u een achterstand, hè? En dan... Alé dan nog kansarmoede, die taalarmoede... We hebben echt kinderen zonder moedertalen. Dat is wel... Kan je dat opvangen als school? Dat weet ik niet. (Respondent 4)

Zowel bij de leerlingen met een hoge SES als bij leerlingen met een lage SES vertelden een aantal respondenten over specifieke onderwijsbehoeften bij de leerlingen waarover ze twijfelden. Voorbeelden hiervan waren leerstoornissen zoals dyslexie en dyscalculie.

4.2.2. Twijfels over het getuigschrift

De meeste respondenten gaven aan dat de twijfels over het al dan niet toekennen van het getuigschrift begonnen rond het 4^e of 5^e leerjaar. Dan wordt de leerstof doorgaans wat moeilijker en wordt de leerachterstand volgens de respondenten beter zichtbaar. Vooral wiskunde werd door de meeste respondenten genoemd als het vak waar de leerlingen het moeilijk mee hadden. De respondenten vertelden dat de eerste moeilijkheden wel al eerder opgemerkt werden, soms zelfs al in de kleuterschool.

Bij de meerderheid van de respondenten waren de twijfels vooral gebaseerd op de eigen toetsen en observaties. Ze gaven aan dat de leerlingen minder mee waren in de klas, de leerstof moeilijker konden vasthouden en dat schoolresultaten achteruit gingen. Respondent 1 stelde bijvoorbeeld: “Maar het viel op, vooral van wiskunde dan, dat zij de dingen niet vasthield dus optellen of optellen tot 20. Ja, dat die automatisatie, de maaltafels. Die kreeg zij niet onder de knie”. De meeste respondenten vertelden dat ze merkten dat de leerlingen moeite hadden met wiskundig inzicht en begrip, en dat ze lager scoorden op toetsen van grotere gehelen.

Alle respondenten gaven aan redelijke aanpassingen te hebben gedaan voor de leerlingen bij wie ze twijfelden over het getuigschrift. Onder andere aangepaste toetsen, onthoudkaarten, rekenmachines

en voorleessoftware werden door de respondenten genoemd. De meeste respondenten vertelden dat de leerlingen tijdens zowel de eigen als de gestandaardiseerde toetsen gebruik mochten maken van dezelfde redelijke aanpassingen als in de les.

Voor een aantal respondenten was het dalende welbevinden van de leerling een alarmbel. Ze zagen dat leerlingen faalangstig werden of andere sociaal-emotionele problemen vertoonden.

Bij een minderheid van de respondenten waren slechte resultaten op de leerlingvolgsysteem-toetsen (LVS), bijvoorbeeld scores in de E-zone, een reden tot twijfel over het getuigschrift. Enkele respondenten gaven hierbij voorbeelden van leerlingen die op de LVS-toetsen een leesniveau van enkele klassen lager hadden dan hun eigenlijke leerjaar.

Naast de schoolinterne aanpassingen, gaf de meerderheid van de respondenten aan dat er ook naar externe ondersteuning werd gezocht. De meeste respondenten vertelden dat de ouders betrokken waren in het traject en er alles aan deden om hun kind te helpen. Enkele respondenten merkten hierbij op dat ouders met een lagere SES soms niet over de financiële middelen beschikten om bijvoorbeeld naar een logopedist te gaan.

4.2.3. Beslissing over het getuigschrift

Alle respondenten vertelden dat er voor het al dan niet toekennen van het getuigschrift meestal gekeken werd naar de resultaten op de gewone toetsen en de eigen observaties in de klas. De meerderheid van de respondenten gaf hierbij aan dat er bij twijfelgevallen specifiek gekeken werd naar in hoeverre de leerling de minimumdoelstellingen behaald had. Aangezien deze doelen vaak lager lagen dan wat er in klas verwacht werd, kregen leerlingen met slechte resultaten die wel de minimumdoelen behaalden, toch hun getuigschrift. Een paar respondenten gaven aan de resultaten van de gestandaardiseerde toetsen mee te nemen in hun beslissing. Ze legden deze dan naast de eigen toetsen of kijken waar de leerling zich situeert tegenover de normgroep.

De meerderheid van de respondenten gaf aan niet enkel naar de punten te kijken voor de beslissing over het getuigschrift. Sommige respondenten benoemden dat ze naar het totale plaatje van de leerling keken en dat de evaluatie gebeurde op basis van het algemene beeld.

De meerderheid van de respondenten probeerde op basis van hun toekomstperspectief voor de leerling in te schatten of ze deze het best in 1A of 1B zagen functioneren. Een aantal respondenten noemden het welbevinden van de leerling als belangrijk argument. Zij gaven aan dat voor faalangstige kinderen de rust van een 1B soms te verkiezen valt boven de stress van een 1A. Een goede werkhouding en sterke motivatie werden door meerdere respondenten gezien als zaken die de slaagkansen van een leerling in 1A kunnen verhogen.

De meerderheid van de respondenten gaf aan ook rekening te houden met de voorkeuren van de leerling zelf. Vooral bij kinderen die zelf aangeven interesse te hebben in een richting uit de B-stroom, werd dit als argument meegenomen om het getuigschrift bij twijfel toch niet toe te kennen. De meeste respondenten vertelden dat ze bij de eindbeslissing weinig tot geen rekening houden met de voorkeur van de ouders.

Een kind waarbij dat je twijfelt getuigschrift of niet, en die weet eigenlijk al dat die in een bouwsector terecht wil komen of in de verzorging ofzo, dan kan je al vlugger oriënteren toch naar een B-stroom dan naar een A-stroom. (Respondent 3)

In de aanloop naar de beslissing spelen de ouders volgens alle respondenten een grotere rol. Alle respondenten gaven aan dat de twijfels en de voorgestelde aanpak in de loop van het traject met de ouders besproken werden. De meeste respondenten gaven aan hier verschillende ervaringen mee te hebben. Soms waren de ouders volledig mee in het verhaal. De respondenten vertelden dat ze dan constructief met hen in gesprek konden gaan over bijvoorbeeld aangepaste trajecten. In sommige gevallen was er weerstand om bijvoorbeeld sommige doelen los te laten en aangepaste toetsen te geven. Deze ouders wilden dat hun kind gewoon dezelfde toetsen bleef doen als de rest van de klas. Enkele respondenten merkten hierbij een verschil tussen ouders met een hoge en een lage SES.

Zeker bij kinderen die uit de middenklasse of hogere sociale klasse [komen] ‘beroeps’ benoemen, dat is moeilijk . . . ‘Daar hoort mijn kind niet thuis’ zo dat... Dat vind ik het grote verschil dat dat veel moeilijker bespreekbaar is bij dan bij onze SES-kinderen. (Respondent 1)

4.2.4. Voor- en nadelen van de verschillende informatiebronnen

De meerderheid van de respondenten gaf aan dat ze de netgebonden toetsen vooral op schoolniveau gebruiken, en veel minder op leerlingniveau. Wat de evaluatie van de leerlingen betreft, vertelden de meeste respondenten dat ze op basis van hun eigen toetsen en observaties doorgaans een goed beeld hebben over de capaciteiten van een leerling. Sommige respondenten gaven echter aan dat de vragen in de methodetoetsen vaak erg lijken op de vragen uit de methode zelf. Leerlingen konden hier dan gericht op oefenen, wat het voor de leerkracht moeilijker maakte om na te gaan of de leerling de leerstof echt begreep. Sommige respondenten vonden het daarom handig om de resultaten van een neutrale, niet-methodegebonden toets zoals de IDP of LVS ernaast te leggen als vergelijking.

De methodetoetsen zijn eigenlijk wel wat gelinkt aan de manier van vragen stellen in de werkboek. Ja, dan heb ik ook kinderen die eigenlijk zoveel van die oefeningen samen met mama en papa opnieuw maken, dat die bijna papegaaierwerk doen op een toets maar geen inzicht hebben in wat doe ik hier eigenlijk? (Respondent 3)

Daarnaast gaven meerdere respondenten aan dat de methodetoetsen vaak verder gaan dan de minimumdoelen. De meerderheid van de respondenten legt daarom bij twijfelgevallen hun toetsen naast de eindtermen of leerplandoelen. Een aantal van hen gaf echter aan dat dit wel wat zoekwerk vergt. Een minderheid van de respondenten vond het daarom gemakkelijk dat de IDP's op de leerplannen gericht zijn, die nauwer aansluiten bij de eindtermen dan dat de inhouden van de methodes doen.

Een ander voordeel van de netgebonden toetsen dat door een aantal respondenten genoemd werd, is dat de IDP's en LVS-toetsen kunnen helpen om door een andere bril naar de evaluatie van de leerling te kijken. Sommige respondenten gebruikten de IDP's en LVS-toetsen ook om de kwaliteit van de eigen toetsen na te gaan. Zolang de resultaten in dezelfde lijn liggen is het oké volgens deze respondenten en als er verschillen zijn, is het nodig om na te gaan hoe dit komt. De meeste respondenten gaven aan dat de resultaten doorgaans overeenkomen. Doordat de resultaten meestal in dezelfde lijn liggen, twijfelden een aantal respondenten echter aan de meerwaarde van de netgebonden toetsen ten opzichte van de eigen toetsen.

Want de leerkrachten zeuren daar ook over, ze vinden het een zeer grote tijdsinvestering om dan uiteindelijk geen ander beeld te zien van een kind. Op zich is dat positief want dat toont dat onze toetsen die we in de klas afnemen, de methodetoetsen of de eigen toetsen van de leerkrachten neutraal en goed zijn. Maar dat zorgt er ook voor dat je LVS'en het gevoel geven

van overbodig te zijn. Dus ik vind niet dat dat een ander beeld geeft, maar het geeft mij wel het gevoel het is neutraal. (Respondent 3)

Respondenten die aangaven eerder holistisch te evalueren, vonden dat de IDP's te veel een momentopname zijn, te hard focussen op kennis en te weinig ruimte laten voor differentiatie. Verder gaven de meeste respondenten aan dat de breedte van de onderdelen van de IDP's te beperkt is. Je neemt als school drie vakgebieden af, wat te weinig is voor een volledig beeld over de capaciteiten van de leerlingen. Sommige respondenten merkten op dat ook de IDP's verder durven gaan dan de minimumdoelen. Hierdoor vonden ze deze ook niet volledig betrouwbaar om zich op te baseren.

Ten slotte vermeldden de respondenten een aantal praktische bezwaren zoals de laattijdige beschikbaarheid van de resultaten en het niet vertrouwd zijn met online afnames.

4.2.5. Adviezen bij de beslissing

Het al dan niet toekennen van een getuigschrift resulteert in een oriëntering naar 1A of 1B in het secundair onderwijs. De meeste respondenten gaven aan dat ze bij dubio meestal het voordeel van de twijfel geven. Op die manier wilden ze nog alle opties voor de leerling openhouden, tenzij ze dachten dat dit een negatieve invloed zou hebben op het welbevinden. Sommige respondenten vertelden dat ze het getuigschrift misschien te snel geven.

Daarnaast gaven een paar respondenten aan dat ze de ouders een advies geven over welke school hen het meest geschikt lijkt voor de leerling. Deze respondenten vertelden dat ze een goed zicht hebben op de scholen in de buurt en weten welke scholen bijvoorbeeld veel zorg of veel structuur bieden. Ten slotte gaven sommige respondenten ook een advies over een geschikte studierichting voor de leerling. Bij leerlingen die het getuigschrift maar nipt behaalden, werd bijvoorbeeld het advies gegeven om te kiezen voor een richting met minder wiskunde.

Ouders doen hun goesting met het advies van onze leerkrachten natuurlijk, maar we merken wel dat als zij zegt van jouw kind zou... Alé die zou echt die of die richting kunnen doen hé dat is een goede richting voor haar, maar ik zou dat dan wel in die of die school doen.
(Respondent 2)

5. Discussie en conclusie

Dit onderzoek bracht aan de hand van een survey en interviews in kaart op welke manier het gebruik van netgebonden toetsen bij de beslissing over het toekennen van een getuigschrift basisonderwijs, bijdraagt tot gelijkere onderwijskansen. Bij de eerste onderzoeksvraag lag de focus enerzijds op hoe betrouwbaar en valide de leraren de informatie vonden die ze ter beschikking hadden, en anderzijds op de mate waarin dit een invloed heeft op hun vooroordelen tegenover GOK-leerlingen. De tweede onderzoeksvraag ging dieper in op het beslissingsproces in het algemeen, argumenten om informatiebronnen wel of niet te gebruiken en impliciete biases die hiermee gepaard kunnen gaan.

In deze sectie worden de bevindingen uit beide onderzoeksvragen geïntegreerd. We gaan daarbij in op de voornaamste bevindingen die uit dit onderzoek naar voren komen en koppelen deze aan de literatuur. Vervolgens worden er enkele beperkingen van dit onderzoek benoemd, en doen we suggesties voor vervolgonderzoek.

5.1. Bevindingen uit het huidige onderzoek

Algemeen kunnen we uit het huidige onderzoek concluderen dat de klassenraad maar in beperkte mate gebruik maakt van de netgebonden toetsen bij de beslissing over het al dan niet toekennen van het getuigschrift basisonderwijs. De informatie uit deze toetsen draagt maar weinig bij tot het rationele beslissingsproces en heeft weinig invloed op de gelijke onderwijskansen. In het huidige onderzoek komt naar voren dat er in het intuïtieve beslissingsproces enkele valkuilen zijn met betrekking tot de gelijke onderwijskansen.

5.1.1. Rationele en intuïtieve processen

In lijn met de wetenschappelijke literatuur toont ook het huidige onderzoek aan dat leraren rationele en intuïtieve beslissingsprocessen combineren bij het nemen van beslissingen over het getuigschrift basisonderwijs. De leraren verzamelen data aan de hand van verschillende soorten toetsen. Ze baseren zich vooral op de eigen toetsen, aangevuld met gestandaardiseerde toetsen. De argumenten die de leraren aanhalen contra de netgebonden toetsen, zoals het onvolledige beeld dat ze geven over de capaciteiten van de leerling, liggen in lijn met eerder onderzoek (Allal, 2013; Brookhart, 2013). Ook de redenen om de netgebonden toetsen wel als bijkomende informatiebron te gebruiken, zoals de neutrale bril die ze kunnen bieden, komen overeen met bevindingen uit de literatuur (Brookhart, 2013; Reskin, 2003).

Opvallend is dat leraren deze combinatie meer betrouwbaar vinden dan leraren die zich voornamelijk baseren op de netgebonden toetsen. Een verklaring kan liggen in het feit dat leraren bij twijfel de eigen toetsen naast de eindtermen leggen bij de afweging over het getuigschrift. Dit geeft hen mogelijk het gevoel over een objectieve standaard te beschikken. Ook het feit dat leraren aangeven dat de resultaten op de netgebonden toetsen bijna altijd in lijn liggen met de resultaten op de eigen toetsen, kan maken dat leerkrachten veel vertrouwen hebben in hun eigen toetsen.

De leraren geven echter aan dat het niet gemakkelijk is om de eindtermen uit de toetsen te halen, en dat de lat voor het al dan niet behalen van de eindtermen vaag is. Dit maakt dat leraren voor het al dan niet toekennen van het getuigschrift, ook met rationeel verzamelde data, in grote mate beroep doen op hun intuïtie wanneer ze de eindbeslissing nemen. Deze bevinding ligt in lijn met eerder onderzoek (Vanlommel, 2018; Vanlommel et al., 2021). Een mogelijke valkuil hierbij is dat leraren op basis van dezelfde informatie toch een verschillende beslissing zouden nemen (Allal, 2013).

Daarnaast maken leraren bij het nemen van de beslissing ook gebruik van intuïtief verzamelde data

zoals de motivatie en interesse van de leerling. Ook uitspraken als ‘naar het algemene plaatje te kijken’ zijn voornamelijk intuïtief.

5.1.2. Biases tegenover GOK-leerlingen

In het huidige onderzoek liggen de resultaten op de schaal met betrekking tot de negatieve percepties lager dan in het oorspronkelijke onderzoek van Lorenz (2021). Ook in de interviews kwamen er weinig tot geen stereotype uitspraken over GOK-leerlingen aan bod. Voor het huidige onderzoek kunnen we dus concluderen dat er geen indicaties zijn van expliciete vooroordelen tegenover GOK-leerlingen bij het al dan niet toekennen van het getuigschrift. Noch de mate van gebruik van de netgebonden toetsen, noch het gevoel van betrouwbaarheid en validiteit van de informatie had een invloed op deze negatieve percepties. Deze bevindingen wijken af van eerder onderzoek dat stelt dat gestandaardiseerde toetsen bijdragen aan verminderde stereotypen tegenover leerlingen met een lagere SES (Bol et al., 2014; Reskin, 2003; van de Werfhorst, 2015).

Daarnaast kwamen er in de huidige studie enkele zaken naar voren die volgens de literatuur kunnen bijdragen tot het benadelen van leerlingen met een lagere SES. Zo gaven leraren in de interviews aan te kijken naar de motivatie en werkhouding van de leerling wanneer ze hun inschatting maken voor de oriëntering. Dit zijn vaardigheden die bij leerlingen met lagere SES doorgaans lager worden ingeschat (Boone & Van Houtte, 2013; Dunkake & Schuchart; 2015). Ook het belang dat gehecht wordt aan de voorkeur van de leerling kan leerlingen met een lagere SES benadelen. Leerlingen met een lagere SES zijn via hun ouders doorgaans meer vertrouwd met arbeidsgerichte richtingen. Dit vergroot de kans dat ze hier interesse in tonen, ongeacht hun capaciteiten (Jæger, 2011).

In het huidige onderzoek komt ook de rol van de ouders sterk naar voren. Ten eerste hebben ouders met een hogere SES meer middelen om externe ondersteuning in te schakelen zoals logopedie. Dit geeft hen een voordeel tegenover ouders met een lagere SES, die deze middelen vaak niet hebben. Deze bevindingen liggen in lijn met het onderzoek van Geven et al. (2018). Ten tweede blijkt uit het huidige onderzoek dat ouders met een hogere SES in de loop van het traject langer blijven aandringen bij de school om de kansen op een getuigschrift open te houden voor hun kind, terwijl ouders met een lagere SES sneller meegaan in het advies van de school. Ook eerder onderzoek benoemde het voordeel dat ouders met hogere SES hebben wanneer het gaat over de communicatie met de school (Barg, 2012; Lareau, 2015). In tegenstelling tot onderzoek van Barg (2012) geven de leraren in het huidige onderzoek aan bij de eindbeslissing weinig rekening te houden met de voorkeur van de ouders.

5.1.3. Implicaties voor de praktijk

Het beperkte gebruik van de netgebonden toetsen betekent niet dat leraren per definitie weigerachtig staan tegenover het gebruik van gestandaardiseerde toetsen. Uit het huidige onderzoek blijkt eerder dat de netgebonden toetsen de rol van objectieve lat volgens de leraren onvoldoende kunnen vervullen. Dit is niet verwonderlijk, aangezien de netgebonden toetsen in de eerste plaats bedoeld zijn om te gebruiken op schoolniveau (Katholiek Onderwijs Vlaanderen, 2021b; OVSG, 2015; Penninckx et al., 2017). Het gevoel over weinig objectieve informatie te beschikken, maakt dat leraren meer beroep doen op hun intuïtie (Geven et al., 2018). Hierbij bestaat het risico dat de beslissing gekleurd wordt door impliciete stereotypen tegenover leerlingen met een lagere SES.

Vanuit de bevindingen uit het huidige onderzoek kunnen we vooruitblikken naar de Vlaamse toetsen die vanaf schooljaar 2023-2024 zullen worden afgenomen. Meer dan dat bij de netgebonden toetsen het geval was, kunnen deze toetsen meegenomen worden in de beoordeling op leerlingniveau (Vlaamse Regering, 2022). Eén van de doelstelling van de Vlaamse toetsen is de sociale gelijkheid

versterken (Steunpunt, 2022; Vlaamse Regering, 2022). Om deze ambitie waar te maken, zullen de Vlaamse toetsen de leraren voldoende informatie moeten geven over de mate waarin de leerling de eindtermen bereikt heeft. Het huidige onderzoek leert ons dat wanneer deze koppeling onvoldoende is, er tijdens de interpretatie van de data veel ruimte blijft voor een subjectief oordeel.

5.2. Beperkingen van het huidige onderzoek

Een beperking van het huidige onderzoek is dat het gebaseerd is op een eerder kleine steekproef. Dit maakt dat er enige terughoudendheid nodig is bij het veralgemenen van de resultaten. Daarnaast maakten alle respondenten uit het kwalitatieve luik gebruik van de IDP's als netgebonden toets. De bevindingen uit de interviews kunnen dus niet zomaar doorgetrokken worden naar gebruikers van de OVSG-toets. Hierbij dient ook opgemerkt te worden dat de IDP's minder op het leerlingniveau gericht zijn dan de OVSG-toets (Katholiek Onderwijs Vlaanderen, 2021c; OVSG, 2015; STEP, 2017).

Een tweede beperking bestaat eruit dat de resultaten uit het kwantitatieve luik gebaseerd zijn op zelfrapportagevragenlijsten. Vooral bij de schaal 'betrouwbaarheid en validiteit van de informatie' is het belangrijk om in het achterhoofd te houden dat het gaat om het gevoel van de respondent. Dit komt mogelijks niet volledig overeen met de effectieve betrouwbaarheid en validiteit. Hetzelfde geldt voor de stereotypen tegenover GOK-leerlingen. Het antwoord op de stellingen komt niet per definitie overeen met het effectieve gedrag dat de respondent zou stellen.

5.3. Suggesties voor vervolgonderzoek

Het verband tussen gestandaardiseerde toetsen en gelijke onderwijskansen kan nog uitgediept worden in verder onderzoek, zowel op kwantitatief als kwalitatief vlak.

In een kwantitatief vervolgonderzoek zou op basis van overheidsdata in kaart gebracht kunnen worden in welke mate de OKI-kenmerken samenhangen met het al dan niet toekennen van het getuigschrift. Vervolgens kan worden nagegaan in welke mate het gebruik van gestandaardiseerde toetsen een invloed heeft op dit verband. Als kwalitatieve piste kan het interessant zijn om enkele klassenraden te observeren wanneer de beslissing wordt genomen. Dit zou onder andere inzage kunnen geven in welke argumenten worden aangehaald en welke infobronnen worden gebruikt bij het nemen van de beslissing.

Door deze aanvullende onderzoekspistes te verkennen, kunnen we niet alleen onze kennis vergroten over de complexe interacties tussen gestandaardiseerde toetsen en gelijke onderwijskansen, maar ook concrete aanbevelingen formuleren om de onderwijspraktijk verder te verbeteren en gelijke onderwijskansen te bevorderen voor alle leerlingen.

6. Referentielijst

- Agirdag, O., Korkmazer, B. (2015). Etnische ongelijkheid in het onderwijs. In Dierckx, D., Coene, J., Raeymaeckers, P., & van der Burg, M. (red.), *Armoede en Sociale Uitsluiting. Jaarboek 2015*. (pp. 231-249). Leuven: Acco. Geraadpleegd van https://www.researchgate.net/publication/287645260_Etnische_ongelijkheid_in_het_onderwijs
- Allal, L. (2013). Teachers' professional judgement in assessment: a cognitive act and a socially situated practice. *Assessment in Education Principles Policy and Practice*, 20(1), 20–34. <https://doi.org/10.1080/0969594x.2012.736364>
- Almeida, F. (2018). Strategies to perform a mixed methods study. *European Journal of Education Studies*, 5(1), 137-151. <https://doi.org/10.5281/ZENODO.1406214>
- Au, W. (2016). Meritocracy 2.0. *Educational Policy (Los Altos, Calif.)*, 30(1), 39–62. <https://doi.org/10.1177/0895904815614916>
- August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice: A Publication of the Division for Learning Disabilities, Council for Exceptional Children*, 20(1), 50–57. <https://doi.org/10.1111/j.1540-5826.2005.00120.x>
- Bakker, J., Denessen, E., & Brus-Laeven, M. (2007). Socio-economic background, parental involvement and teacher perceptions of these in relation to pupil achievement. *Educational Studies*, 33(2), 177–192. <https://doi.org/10.1080/03055690601068345>
- Barg, K. (2013). The influence of students' social background and parental involvement on teachers' school track choices: Reasons and consequences. *European Sociological Review*, 29(3), 565–579. <https://doi.org/10.1093/esr/jcr104>
- Bol, T., Witschge, J., van de Werfhorst, H. G., & Dronkers, J. (2014). Curricular tracking and central examinations: Counterbalancing the impact of social background on student achievement in 36 countries. *Social Forces; a Scientific Medium of Social Study and Interpretation*, 92(4), 1545–1572. <https://doi.org/10.1093/sf/sou003>
- Boone, S. (2013). Onderwijsongelijkheden: de confrontatie van gezinnen met een onderwijssysteem. In Eidhof, B., Van Houtte, M. & Vermeulen, M. (red.), *Sociologen over onderwijs. Inzichten, praktijken en kritieken*. (pp. 219-235). Antwerpen: Garant
- Boone, S., & Van Houtte, M. (2013). Why are teacher recommendations at the transition from primary to secondary education socially biased? A mixed-methods research. *British Journal of Sociology of Education*, 34(1), 20–38. <https://doi.org/10.1080/01425692.2012.704720>
- Brookhart, S. M. (2013). The use of teacher judgement for summative assessment in the USA. *Assessment in Education Principles Policy and Practice*, 20(1), 69–90. <https://doi.org/10.1080/0969594x.2012.703170>
- Broomes, O. P. (2013). More than a new country: Effects of immigration, home language, and school mobility on elementary students' academic achievement over time. *Education Policy Analysis Archives*, 21(48). <https://doi.org/10.14507/epaa.v21n48.2013>
- Calarco, J. M. (2011). "I need help!" social class and children's help-seeking in elementary school. *American Sociological Review*, 76(6), 862–882. <https://doi.org/10.1177/0003122411427177>

- Coburn, C. E., & Turner, E. O. (2011). Research on data use: A framework and analysis. *Measurement: Interdisciplinary Research and Perspectives*, 9(4), 173–206. <https://doi.org/10.1080/15366367.2011.626729>
- De Maeyer, S., van Daal, T., Lesterhuis, M. (2020). *Open Leerpakket Meten en Observeren* [Ongepubliceerde cursustekst]. Universiteit Antwerpen.
- Dunkake, I., & Schuchart, C. (2015). Stereotypes and teacher characteristics as an explanation for the class-specific disciplinary practices of pre-service teachers. *Teaching and Teacher Education*, 50, 56–69. <https://doi.org/10.1016/j.tate.2015.04.005>
- Ferguson, H., Bovaird, S., & Mueller, M. (2007). The impact of poverty on educational outcomes for children. *Paediatrics & Child Health*, 12(8), 701–706. <https://doi.org/10.1093/pch/12.8.701>
- Franck E. & Nicaise I. (2018). *Ongelijkheden in het Vlaamse onderwijssysteem: verbetering in zicht? Een vergelijking tussen PISA 2003 en 2015*. Leuven: HIVA, Gent: SONO, 60p.
- Franck, E., Nicaise, I. & Lavrijsen, J. (2017). *Extra middelen, meer gelijke onderwijskansen? De effectiviteit van de compensatiefinanciering voor scholen met leerlingen uit kansengroepen onder de loep*. Steunpunt Onderwijsonderzoek, KU Leuven
- Geven, S., Batruch, A. & van de Werfhost, H. (2018). *Inequality in Teacher Judgements, Expectations and Track Recommendations: A Review Study*. Amsterdam: Universiteit van Amsterdam
- Glock, S., Krolak-Schwerdt, S., & Pit-ten Cate, I. M. (2015). Are school placement recommendations accurate? The effect of students' ethnicity on teachers' judgments and recognition memory. *European Journal of Psychology of Education*, 30(2), 169–188. <https://doi.org/10.1007/s10212-014-0237-2>
- GO! (2022). *Standpunt getuigschrift basisonderwijs*. Geraadpleegd op 2 januari 2023, van <https://pro.g-o.be/blog/Documents/Standpunt%20getuigschrift%20basisonderwijs.pdf>
- Gunzler, D., Chen, T., Wu, P., & Zhang, H. (2013). Introduction to mediation analysis with structural equation modeling. *Shanghai Archives of Psychiatry*, 25(6), 390–394. <https://doi.org/10.3969/j.issn.1002-0829.2013.06.009>
- Hays, R. D., Hayashi, T., & Stewart, A. L. (1989). A five-item measure of socially desirable response set. *Educational and Psychological Measurement*, 49(3), 629–636. <https://doi.org/10.1177/001316448904900315>
- Jæger, M. M. (2011). Does cultural capital really affect academic achievement? New evidence from combined sibling and panel data. *Sociology of Education*, 84(4), 281–298. <https://doi.org/10.1177/0038040711417010>
- Katholiek Onderwijs Vlaanderen (2021a). *Werken aan gelijke onderwijskansen*. Geraadpleegd op 2 januari 2023, van <https://pro.katholiekonderwijs.vlaanderen/gelijke-onderwijskansen/werken-aan-gelijke-onderwijskansen>
- Katholiek Onderwijs Vlaanderen (2021b). *Klasrapport einde BaO 2020-2021*.
- Katholiek Onderwijs Vlaanderen (2021c). *IDP-stappenplan*. Geraadpleegd op 20 maart 2023, van <https://pro.katholiekonderwijs.vlaanderen/kwaliteitsinstrumenten/idpstappenplan>
- Katholiek Onderwijs Vlaanderen (2022a). *Getuigschrift basisonderwijs*. Geraadpleegd op 20 maart 2023, van <https://pro.katholiekonderwijs.vlaanderen/leerlingen-evalueren-in-het-basisonderwijs/getuigschrift-basisonderwijs>
- Katholiek Onderwijs Vlaanderen (2022b). *IDP-informatie*. Geraadpleegd op 2 januari 2023, van <https://pro.katholiekonderwijs.vlaanderen/evaluatiebox-basisonderwijs/idp-informatie>

- Klenowski, V. (2014). Towards fairer assessment. *Australian Educational Researcher*, 41(4), 445–470. <https://doi.org/10.1007/s13384-013-0132-x>
- Koch, A. J., D’Mello, S. D., & Sackett, P. R. (2015). A meta-analysis of gender stereotypes and bias in experimental simulations of employment decision making. *The Journal of Applied Psychology*, 100(1), 128–161. <https://doi.org/10.1037/a0036734>
- Lareau, A. (2015). Cultural knowledge and social inequality. *American Sociological Review*, 80(1), 1–27. <https://doi.org/10.1177/0003122414565814>
- Leighton, J. P. (2004). Avoiding misconception, misuse, and missed opportunities: The collection of verbal reports in educational achievement testing. *Educational Measurement: Issues and Practice*, 23(4), 6–15. <http://dx.doi.org/10.1111/j.1745-3992.2004.tb00164.x>
- Lopez-Agudo, L. A., González-Betancor, S. M., & Marcenaro-Gutierrez, O. D. (2021). Language at home and academic performance: The case of Spain. *Economic Analysis and Policy*, 69, 16–33. <https://doi.org/10.1016/j.eap.2020.11.003>
- Lorenz, G. (2021). Subtle discrimination: do stereotypes among teachers trigger bias in their expectations and widen ethnic achievement gaps? *Social Psychology of Education: An International Journal*, 24(2), 537–571. <https://doi.org/10.1007/s11218-021-09615-0>
- Maerten-Rivera, J., Myers, N., Lee, O., & Penfield, R. (2010). Student and school predictors of high-stakes assessment in science. *Science Education*, 94(6), 937–962. <https://doi.org/10.1002/sce.20408>
- McMullin, C. (2023). Transcription and qualitative methods: Implications for third sector research. *VOLUNTAS International Journal of Voluntary and Nonprofit Organizations*, 34(1), 140–153. <https://doi.org/10.1007/s11266-021-00400-3>
- Michelmore, K., & Dynarski, S. (2017). The gap within the gap. *AERA Open*, 3(1). <https://doi.org/10.1177/2332858417692958>
- Mijs, J. & van de Werfhorst, H. G. (2013). Het onderwijsstelsel en kansengelijkheid: op zoek naar een meritocratie. In Eidhof, B., Van Houtte, M. & Vermeulen, M. (red.), *Sociologen over onderwijs. Inzichten, praktijken en kritieken*. (pp. 199-217). Antwerpen: Garant
- Nieuwenhuis, J., & Hooimeijer, P. (2016). The association between neighbourhoods and educational achievement, a systematic review and meta-analysis. *Journal of Housing and the Built Environment*, 31(2), 321–347. <https://doi.org/10.1007/s10901-015-9460-7>
- Penninckx, M., Vanhoof, J., Quintelier, A., De Maeyer, S., Van Petegem, P. (2017). *Zicht op leerwinst. Scenario’s voor gestandaardiseerde toetsen*. Leuven: Acco.
- Pit-ten Cate, I. M., Krolak-Schwerdt, S., & Glock, S. (2016). Accuracy of teachers’ tracking decisions: short- and long-term effects of accountability. *European Journal of Psychology of Education*, 31(2), 225–243. <https://doi.org/10.1007/s10212-015-0259-4>
- Prevo, M. J. L., Malda, M., Emmen, R. A. G., Yeniad, N., & Mesman, J. (2015). A context-dependent view on the linguistic interdependence hypothesis: Language use and SES as potential moderators. *Language Learning*, 65(2), 449–469. <https://doi.org/10.1111/lang.12099>
- Reskin, B. F. (2003). Including mechanisms in our models of ascriptive inequality: 2002 presidential address. *American Sociological Review*, 68(1), 1. <https://doi.org/10.2307/3088900>
- Rogers-Ard, R., Knaus, C. B., Epstein, K. K., & Mayfield, K. (2013). Racial diversity sounds nice; Systems transformation? Not so much. *Urban Education*, 48(3), 451–479. <https://doi.org/10.1177/0042085912454441>

- Schildkamp, K., & Ehren, M. (2013). *Data-based decision making in education: Challenges and opportunities*. Dordrecht: Springer
- Schildkamp, K., Poortman, C. L., & Handelzalts, A. (2016). Data teams for school improvement. *School Effectiveness and School Improvement : An International Journal of Research, Policy and Practice*, 27(2), 228–254. <https://doi.org/10.1080/09243453.2015.1056192>
- Steunpunt (2022). *Toetsen voor onderwijsontwikkeling. Krachtlijnen voor de centrale toetsen in Vlaanderen*. Steunpunt Centrale Toetsen in Onderwijs
- Steunpunt Toetsontwikkeling en Peilingsonderzoek (2017). *Validering van IDP en de OVSG-toets: Eindrapport*. Geraadpleegd op 20 maart 2023, van <https://data-onderwijs.vlaanderen.be/documenten/bestand.ashx?id=7760>
- Timmermans, A. C., Kuyper, H., & van der Werf, G. (2015). Accurate, inaccurate, or biased teacher expectations: Do Dutch teachers differ in their expectations at the end of primary education? *The British Journal of Educational Psychology*, 85(4), 459–478. <https://doi.org/10.1111/bjep.12087>
- Valcke, M. & Standaert, R. (2018). *Onderwijsbeleid in Vlaanderen*. Leuven: Acco
- van de Werfhorst, H. G. (2015). De achterkant van autonomie: maatwerk, sturing en ongelijkheid. *Nederlands Tijdschrift voor Onderwijsrecht en Onderwijsbeleid*, December: 286-293.
- Van Gasse, R., Vanhoof, J., Mathieu, P., Van Petegem, P. (2015). *Informatiegebruik door schoolleiders en leerkrachten*. Garant.
- Van Laere, E., Aesaert, K., & van Braak, J. (2014). The Role of Students' Home Language in Science Achievement: A multilevel approach. *International Journal of Science Education*, 36(16), 2772–2794. <https://doi.org/10.1080/09500693.2014.936327>
- van Leest, A., Hornstra, L., van Tartwijk, J., & van de Pol, J. (2021). Test- or judgement-based school track recommendations: Equal opportunities for students with different socio-economic backgrounds? *The British Journal of Educational Psychology*, 91(1), 193–216. <https://doi.org/10.1111/bjep.12356>
- Vanlommel, K. (2018). *Opening the black box of teacher judgement: The interplay of rational and intuitive processes* [Proefschrift]. Universiteit Antwerpen
- Vanlommel, K., Van Gasse, R., Vanhoof, J., & Van Petegem, P. (2018). Teachers' high-stakes decision making. How teaching approaches affect rational and intuitive data collection. *Teaching and Teacher Education*, 71, 108–119. <https://doi.org/10.1016/j.tate.2017.12.011>
- Vanlommel, K., Van Gasse, R., Vanhoof, J., & Van Petegem, P. (2021). Sorting pupils into their next educational track: How strongly do teachers rely on data-based or intuitive processes when they make the transition decision? *Studies in Educational Evaluation*, 69(100865), 100865. <https://doi.org/10.1016/j.stueduc.2020.100865>
- Vlaams ministerie van onderwijs en vorming (2016). *Getuigschrift basisonderwijs in het gewoon lager onderwijs*. Geraadpleegd op 2 januari 2023, van <https://onderwijs.vlaanderen.be/nl/ouders/leren-en-evalueren/getuigschriften-diplomas-en-certificaten/getuigschrift-basisonderwijs/getuigschrift-basisonderwijs-in-het-gewoon-lager-onderwijs>
- Vlaams ministerie van onderwijs en vorming (2017). *Toolkit gevalideerde toetsen basisonderwijs*. Geraadpleegd op 20 maart 2023, van <https://onderwijs.vlaanderen.be/nl/onderwijspersoneel/van-basis-tot->

volwassenenonderwijs/lespraktijk/toetsen-voor-scholen/toolkit-gevalideerde-toetsen-basisonderwijs

- Vlaamse Regering (2022). *Ontwerp van decreet over de Vlaamse toetsen in het onderwijs* [Parlementair document]. Brussel: Vlaams Parlement. Geraadpleegd op 6 april 2023, van <https://docs.vlaamsparlement.be/pfile?id=1927743>
- Wilson, C. (2014). Semi-Structured Interviews. In *Interview Techniques for UX Practitioners* (pp. 23–41). <https://doi.org/10.1016/b978-0-12-410393-1.00002-8>
- Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement, 76*(6), 913–934. <https://doi.org/10.1177/0013164413495237>
- Wyatt-Smith, C. & Klenowski, V. (2013). Explicit, latent and meta-criteria: types of criteria at play in professional judgement practice. *Assessment in Education: Principles, Policy & Practice, 20*(1), 35-52. <https://doi.org/10.1080/0969594X.2012.725030>

7. Bijlagen

7.1. Bijlage 1: Verklaring op eer

7.2. Bijlage 2: Survey

Inleiding

Scholen hebben veel data tot hun beschikking, waaronder de resultaten van gestandaardiseerde toetsen, ontwikkeld binnen een koepel (interdiocesane proeven, OVSG-toets...). Er bestaat in de literatuur nog veel onduidelijkheid over de invloed van gestandaardiseerde testen op de gelijke onderwijskansen van leerlingen. In dit masterproefonderzoek willen we nagaan in hoeverre er een verband is tussen het gebruik van de gestandaardiseerde toetsen en percepties tegenover GOK-leerlingen.

In het kader van dit onderzoek vragen we aan directies, zorgcoördinatoren en leerkrachten die deel uitmaken van de klassenraad basisonderwijs om een korte vragenlijst in te vullen. Deze zal zo'n 5 minuten in beslag nemen en is volledig anoniem. Indien u vrijblijvend interesse wilt aangeven om in een tweede fase deel te nemen aan een interview, kan u aan het einde van de vragenlijst uw contactgegevens achterlaten.

Mocht u nog vragen hebben of meer informatie wensen, kan u ons contacteren via onderstaande gegevens:

Onderzoeker: Stijn Kerkhofs (stijn.kerkhofs@student.uantwerpen.be)

Promotor: dr. Roos Van Gasse (roos.vangasse@uantwerpen.be)

Algemene vragen

1. Wat is uw gender?
 - Man
 - Vrouw
 - Andere
2. Tot welk onderwijsnet behoort de school waar u voor het grootste deel van uw opdracht bent tewerkgesteld?
 - Gemeenschapsonderwijs (GO!)
 - Gesubsidieerd officieel onderwijs (OVSG, POV)
 - Gesubsidieerd vrij onderwijs (Katholiek Onderwijs, OKO,...)
3. Welke categorie omschrijft de functie waar u het meeste tijd aan besteedt het beste?
 - Lid van het directieteam
 - Lid van het zorgteam
 - Klasleerkracht 6^e leerjaar
 - Leerkracht, behalve klasleerkracht 6^e leerjaar (ander leerjaar, lichamelijke opvoeding,...)
 - Andere: *tekstveld*
4. Hoe lang oefent u deze functie reeds uit? Noteer het aantal jaren. * tekstveld met enkel cijfers*

Mate van gebruik

5. In welke mate maakt u gebruik van gestandaardiseerde toetsen, ontwikkeld binnen een koepel (interdiocesane proeven, OVSG-toets...) bij het nemen van beslissingen over het al dan niet toekennen van een getuigschrift basisonderwijs aan een leerling?

- Ik gebruik deze gestandaardiseerde toetsen niet. Ik baseer me enkel op de zelfgemaakte leerlingentoetsen.
- Ik baseer me vooral op de zelfgemaakte leerlingentoetsen. Ik gebruik de gestandaardiseerde toetsen als bijkomende informatiebron.
- Ik baseer me evenveel op de zelfgemaakte leerlingentoetsen als op de gestandaardiseerde toetsen.
- Ik baseer me vooral op de gestandaardiseerde toetsen. Ik gebruik de zelfgemaakte leerlingentoetsen als een bijkomende informatiebron.
- Ik baseer me enkel op de gestandaardiseerde toetsen.

Betrouwbaarheid en validiteit

In de volgende stellingen gaan we dieper in op de afstemming van beschikbare informatie op uw behoeften als gebruiker. In welke mate bent u het eens met de volgende stellingen?

(helemaal oneens – eerder oneens – noch oneens/noch eens – eerder eens – helemaal eens – weet niet/niet van toepassing)

6. De informatie die ik ter beschikking heb, geeft een juist beeld.
7. De informatie die ik ter beschikking heb, is correct.
8. De informatie die ik ter beschikking heb, is betrouwbaar.
9. De informatie die ik ter beschikking heb, is objectief.
10. De informatie die ik ter beschikking heb, meet precies wat wij willen weten.

Percepties tegenover GOK-leerlingen

Volgende stellingen peilen naar enkele percepties tegenover GOK-leerlingen. Onder 'GOK-leerling' verstaan we een leerling die aan minstens één van volgende criteria voldoet: 1) de thuistaal is niet het Nederlands, 2) de moeder heeft geen diploma van minstens hoger secundair onderwijs, 3) de leerling ontvangt een schooltoelage, 4) de leerling woont in een buurt met veel schoolse vertraging. In welke mate bent u het eens met onderstaande stellingen?

(helemaal oneens – eerder oneens – noch oneens/noch eens – eerder eens – helemaal eens)

11. GOK-leerlingen zijn gemiddeld minder geïnteresseerd in school-gerelateerde onderwerpen dan andere leerlingen.
12. GOK-leerlingen volgen de lessen gemiddeld minder zorgvuldig dan andere leerlingen.
13. GOK-leerlingen zijn gemiddeld minder gemotiveerd om te leren dan andere leerlingen.
14. GOK-leerlingen spannen zich gemiddeld minder in op school dan andere leerlingen.
15. GOK-leerlingen hebben gemiddeld minder voorkennis dan andere leerlingen.

Relaties met anderen (sociaal wenselijkheid)

Hieronder staan enkele uitspraken over uw relaties met anderen. In hoeverre is elke uitspraak waar of onwaar voor u?

(helemaal oneens – eerder oneens – noch oneens/noch eens – eerder eens – helemaal eens)

16. Ik ben altijd beleefd, zelfs tegen mensen die onvriendelijk zijn.
17. Ik heb wel eens misbruik van iemand gemaakt.
18. Ik probeer soms wraak te nemen in plaats van te vergeven en te vergeten.
19. Ik voel me soms kwaad als ik mijn zin niet krijg.

20. Met wie ik ook praat, ik kan altijd goed luisteren.

Bereidheid voor interview

In een ander luik van het onderzoek gaan we dieper in op hoe leerkrachten basisonderwijs gebruik maken van de resultaten van netgebonden toetsen op leerlingniveau bij het nemen van beslissingen. Hiervoor zoeken we enkele leerkrachten die bereid zijn om deel te nemen aan een (online) interview. Zouden we u hiervoor mogen contacteren? Indien u 'ja' aangeeft, kan u op het volgende scherm uw contactgegevens achterlaten.

21. Bent u bereid om deel te nemen aan een interview?

- Ja
- Nee

Indien ja

22. Wat is uw voornaam? *tekstveld*

23. Wat is uw familienaam? *tekstveld*

24. Wat is uw e-mailadres? *tekstveld*

7.3. Bijlage 3: Interviewleidraad

Kenmerken:

- Semigestructureerd interview
- Duurtijd: +- 45 minuten
- Respondenten: leden van de klassenraad BaO die beslist over het getuigschrift
- Onderwerp: gebruik van netgebonden toetsen bij toekenning getuigschrift
- Onderzoeksvraag: Op welke manier maakt de klassenraad basisonderwijs gebruik van de resultaten van netgebonden toetsen op leerlingniveau bij de toekenning van het getuigschrift?

Onderdeel/vragen	Timing
<p>Introductie</p> <p>Bedankt om deel te nemen aan dit interview. Dit interview kadert binnen de masterproef voor de opleiding opleidings- en onderwijswetenschappen en zal ongeveer 45 minuten duren. Het onderwerp van het interview is hoe netgebonden toetsen gebruikt worden bij de beslissing over een getuigschrift. De resultaten van dit interview zullen anoniem verwerkt worden aan de hand van pseudoniemen voor zowel jezelf als de leerlingen. Mag ik deze videocall opnemen om de gegevens later te verwerken? Graag jouw toestemming voor deze opname. Heb jij nog vragen vooraleer we van start gaan?</p>	2'
<p>Openingsvragen</p> <ol style="list-style-type: none"> 1. Kan je jezelf even voorstellen? <ul style="list-style-type: none"> ○ <i>In welke school werk je? Wat is je functie?</i> 2. Welke netgebonden toets gebruiken jullie op school? 3. Waarvoor gebruiken jullie netgebonden toetsen op school? 	4'
<p>Toelichting verloop</p> <p>Ik bespreek vandaag graag met jou twee casussen waarbij er vorig schooljaar twijfel was over het al dan niet toekennen van het getuigschrift. Liefst één leerling waarvan je denkt dat deze een hogere sociaal economische status en één waarvan je denkt dat deze een lagere sociaal economische status heeft. Met welke leerling zou je graag beginnen?</p>	1'
<p>Inleidende vragen per case</p> <ol style="list-style-type: none"> 4. Kan je X even voorstellen aan mij? <ul style="list-style-type: none"> ○ <i>SES achtergrond bevragen</i> 5. Wanneer begonnen de twijfels of X wel een getuigschrift zou halen? 6. Wat was het precies dat je deed twijfelen? 	6' (x2)
<p>Overgangsvragen per case</p> <ol style="list-style-type: none"> 7. Hoe waren de resultaten van X op de [netgebonden toets], in vergelijking met het beeld dat je had van X? <ul style="list-style-type: none"> ○ <i>Eigen toetsen, andere informatie</i> 	3' (x2)
<p>Kernvragen per case</p> <ol style="list-style-type: none"> 8. Op welke manier heb je de resultaten van de [netgebonden toets] en de andere informatie gebruikt om tot een beslissing te komen over het getuigschrift? 9. Wat is de uiteindelijke beslissing geworden over het getuigschrift? 10. Wat waren de doorslaggevende argumenten in de beslissing? 	8' (x2)

Slot De bedoeling van dit interview was om zicht te krijgen op hoe netgebonden toetsen gebruikt worden bij de beslissing over een getuigschrift. Zijn er zaken die in het interview niet aan bod zijn gekomen waar je zelf toch nog graag iets over wilt vertellen?	2'
---	----

7.4. Bijlage 4: Voorbeeld van een transcript

7.5. Bijlage 5: Codeboom

Name	Files	References
Achtergrondkenmerken	5	20
Andere	5	5
Anderstalig	3	8
Armoede	2	2
HogeSES	2	3
OpleidingMoeder	1	2
Advies	4	23
AofB	4	11
School	2	6
Studierichting	2	6
Eindbeslissing	5	66
AlgemeenBeeld	3	4
DoelenBereikt	4	8
EigenToetsen	5	11
GestandaardiseerdeToetsen	2	3
LInKenmerken	5	16
Gedrag	0	0
Motivatie	1	4
Welbevinden	3	7
Werkhouding	4	5
Toekomstperspectief	4	8
Trajectmatig	2	4
VoorkeurKind	4	6

Name	Files	References
VraagOuders	3	6
GebruikInfobronnen	5	69
EigenToetsen	5	15
NaastEindtermen	4	7
NadelenEigenToetsen	3	7
VoordelenEigenToetsen	1	1
GestandaardiseerdeToetsen	5	54
MateVanGebruik	5	19
NadelenGestandaardiseerdeToetsen	5	21
VoordelenGestandaardiseerdeToetsen	5	14
RedenTwijfel	5	21
EigenToetsEnObs	3	9
Evolutie	1	1
GestandaardiseerdeToets	1	3
LInKenmerken	5	8
Andere	1	1
Gedrag	0	0
Motivatie	1	2
Hoog	1	1
Welbevinden	3	4
Werkhouding	1	1
ResultaatGestandaardiseerdeToetsen	4	12
KomtOvereen	4	8
NietOvereen	3	4

Name	Files	References
VolgenEigenToetsen	1	1
VolgenGestandaardiseerdeToetsen	1	1
Traject	5	70
Aanpak	5	21
Evolutie	4	5
RedelijkeAanpassingen	5	15
Ouders	5	34
StartTwijfel	5	15
VlaamseToetsen	3	7
Negatief	1	2
Postief	2	2